

AD-A219 048

RADC-TR-89-280
Final Technical Report
December 1989

DTIC FILE COPY



4

MULTIPLE BEAM OPTICAL PROCESSING

University of Arizona

DTIC
SELECTE
MAR 14 1990
S D D

Sponsored by
Defense Advanced Research Projects Agency
ARPA Order 5796

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

90 03 13 070

This report has been reviewed by the RADC Public Affairs Division (PA) and is releasable to the National Technical Information Services (NTIS) At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-89-280 has been reviewed and is approved for publication.

APPROVED:



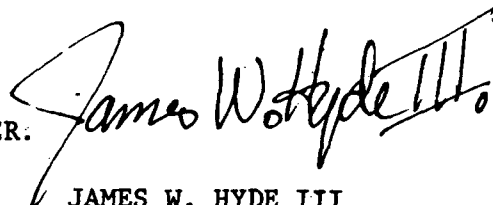
DOROTHY J. JACKSON, 1Lt, USAF
Project Engineer

APPROVED:



JOHN A. GRANIERO
Technical Director
Directorate of Communications

FOR THE COMMANDER:



JAMES W. HYDE III
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (DCLW) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

MULTIPLE BEAM OPTICAL PROCESSING

H. Gibbs
N. Peyghambarian
G. Khitrova
S. Koch
R. Kostuk
A. Macleod
K. Wagner
D. Roude
M. Warren

Contractor: University of Arizona
Contract Number: F30602-87-C-0009
Effective Date of Contract: 29 Jan 87
Contract Expiration Date: 29 Jan 89
Short Title of Work: Multiple Beam Optical Processing
Program Code Number: 7D10
Program Element: 61101E
ARPA Order Number: 5796
Period of Work Covered: Jan 87 - Jan 89

Principal Investigator: Hyatt M. Gibbs
Phone: (602) 621-2941

RADC Project Engineer: Dorothy J. Jackson, 1Lt, USAF
Phone: (315) 330-4092

Approved for public release, distribution unlimited.

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by 1Lt Dorothy J. Jackson, RADC (DCLW) Griffiss AFB NY 13441-5700 under Contract F30602-87-C-0009.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188		
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS N/A			
2a. SECURITY CLASSIFICATION AUTHORITY N/A			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.			
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A						
4. PERFORMING ORGANIZATION REPORT NUMBER(S) N/A			5. MONITORING ORGANIZATION REPORT NUMBER(S) RADC-TR-89-280			
6a. NAME OF PERFORMING ORGANIZATION University of Arizona		6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Prime Air Development Center (DCLW)			
6c. ADDRESS (City, State, and ZIP Code) Optical Sciences Center Tucson AZ 85721			7b. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Advanced Research Projects Agency		8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F30602-87-C-0009			
8c. ADDRESS (City, State, and ZIP Code) 1400 Wilson Boulevard Arlington VA 22209			10. SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
			62702F	E796	01	01
11. TITLE (Include Security Classification) MULTIPLE BEAM OPTICAL PROCESSING						
12. PERSONAL AUTHOR(S) H. Gibbs, N. Peyghambarian, G. Khitrova, S. Koch, R. Kostuk, A. Macleod, K. Wagner, D. Rouede, M. Warren						
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM Jan 87 TO Jan 89		14. DATE OF REPORT (Year, Month, Day) December 1989		15. PAGE COUNT 246
16. SUPPLEMENTARY NOTATION N/A						
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)			
FIELD	GROUP	SUB-GROUP				
20	06	01				
			Optical Bistability Optical Logic Fabry Perot Etalons Holographic Optical Elements Optical Nonlinear Directional Coupler (Continued on reverse)			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The objective of this effort was to develop optical processing technology required to implement high performance signal processing systems. The work concentrated on the evaluation and characterization of GaAs nonlinear etalons and methods of using these etalons in optical processing functions. As a means of determining the potential of etalons for use in optical processing functions several requirements were studied. The optical nonlinearities of the etalons were evaluated for their ability to provide differential gain for cascading and fanout. The optical nonlinearities and gain were characterized as a function of well width in multiple quantum well structures, the carrier lifetime of the nonlinear optical material, and the pulse length and angle of incidence of the switching beam. Methods of providing fan-in of input signals to the etalons were also studied. Methods of providing interconnects to the etalons were developed. These included the use of optical fibers and holographic (Continued on reverse)						
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS				21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL Dorothy J. Jackson, 1Lt, USAF				22b. TELEPHONE (Include Area Code) (315) 330-4092		22c. OFFICE SYMBOL RADC (DCLW)

DD Form 1473, JUN 86

Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

UNCLASSIFIED

Block 18 Continued: Multiple Quantum Wells Photorefractive Crystals
 ZnS and ZnSe Interference Filters Symbolic Substitution
 Associative Memory Pattern Recognition

Block 19 Continued:

optical elements. Several optical processing functions were implemented and the potential of nonlinear etalons for use in these functions was evaluated. The experimental work was performed using ZnSe interference filters due to their relatively inexpensive cost and availability. However, the results and conclusions of the experiments may be applied to the use of GaAs nonlinear etalons as well. These experiments included the implementation of a one bit adder using symbolic substitution, a pattern recognizer that can distinguish between simple three spot patterns, an associative memory which can discriminate against two different input objects, and an all optical compare and exchange module which can compare the magnitude of two digital input signals and make an appropriate exchange decision. These experiments, although rudimentary, may be extendable to larger scales upon further development of the optical processing technology.



Approved	
NTIS	✓
DTIC	✓
Document	
Index	
By	
Date	
Author	
Editor	
Dist	
A-1	

UNCLASSIFIED

SUMMARY

Part A: GaAs Device Research

The concept of *nonlinear* optical decisionmaking is introduced (Section 1). The properties of a nonlinear optical system depend on the strength or intensity of the light incident upon it. Under special conditions, a nonlinear system can exhibit a threshold characteristic. Consequently, an optical transistor can be made to perform logic operations analogous to those of an electronic transistor.

The operation of zinc sulfide (ZnS) interference filters (IFs) is based on thermal, nonlinear optical refraction. ZnS IFs are slow and power-hungry, but they operate in the visible light range with many-watt argon lasers. These properties, we decided, warranted their development, to enable us to perform optical signal processing demonstrations (Part C: Sections 13-15).

Gallium arsenide (GaAs) etalons (Section 4) are particularly attractive for application in nonlinear optical systems. Nonlinear absorption spectra of bulk GaAs, and of 299-, 152-, and 76-Å GaAs/AlGaAs multiple-quantum-well (MQW) samples are reported. Nonlinear index spectra are calculated by Kramers-Kronig transformations of the absorption spectra. It is concluded that the principal nonlinear refractive mechanisms for optical bistability in bulk GaAs etalons are band filling and reduction of the Coulomb enhancement of continuum states. Exciton saturation and bandgap renormalization occur at lower carrier densities, but their contributions are of opposite sign and almost cancel. Exciton saturation nonlinearities dominate for bistability in MQW samples with wells of < 100 Å. Fabrication and testing of GaAs etalons are reported in Section 4, with particular emphasis on uniformity. Molecular beam epitaxy (MBE) seems to be the best method for fabricating reflectors of high uniformity.

Several of the properties necessary for GaAs etalons to be used as all-optical gates in a practical system are demonstrated:

- 1) Cascading of GaAs etalons used as logic gates (Section 5);
- 2) Differential gain, with a change in output 4 times greater than the change in input, with good contrast between high and low states (Section 5);
- 3) Effective fan-in of signals to a GaAs etalon (Section 8).

A theoretical analysis is included, which shows that differential gain is possible only if the pulse lengths are greater than about 10 times the carrier lifetime in the nonlinear optical material (Section 6). In bulk GaAs, the carrier lifetime is on the

order of 3 nanoseconds; the carrier lifetime must therefore be reduced to achieve cascading of GaAs etalons on a picosecond time scale with single-wavelength operation.

The dependence of differential gain in GaAs bistable etalons on the incident angle of the switching beam is studied theoretically for the case of pulsed, single-wavelength operation. Both dispersive and absorptive aspects of the etalon are included. Simulations predict that differential-gain characteristics can be improved significantly over the normal-incidence case by utilizing angular separation of the pump and switching beams. The limitations of the plane-wave approximation and the potential effect of beam walk-off are discussed.

Optical interconnections to etalons are addressed in Sections 8 through 11. A fan-in of two signal beams and one holding beam to achieve AND-gate operation is demonstrated, using bulk-optics polarization beam splitters (Section 8). In addition, GaAs etalons can be used effectively with optical fibers by direct coupling, with no intervening optics (Section 9). Fabrication and testing of holographic optical elements (HOEs) are described in considerable detail in Section 10. A single HOE was made to replace the combination of a computer-generated hologram and lens in forming an array of 1×8 spots; the HOE was used to switch bistability on a ZnS IF. Preliminary results of an experiment designed to show that GaAs etalons may be used with a HOE that produces an array of focused spots comprises Section 11.

Part B: Comparison of Devices for Optical Signal Processing

Here we take a critical look at device characteristics, using the Statement of Work as a guide, and describing the progress achieved and the problems encountered. The difficulty of cascading at high speeds was discovered under this contract; the solution to this problem has not yet been found (Section 12).

Part C: Optical Signal Processing Using ZnS Interference Filters for Nonlinear Decisionmaking

One-bit addition by symbolic substitution is demonstrated by implementation of each of the optical channels required (Section 13). Simple patterns (V and Γ) consisting of three spots are recognized by dividing, shifting, and recombining beams onto bistable ZnS interference filters. This experiment demonstrates AND-gate operation, cascading, and a moderate amount of parallelism, but a laser power of several watts is required and response times are several milliseconds.

An associative memory for fingerprint identification has been constructed using a VanderLugt correlator and an interference filter as a reflective thresholding device.

All-optical compare-and-exchange is demonstrated in ZnS bistable optical devices. The compare-and-exchange demonstration utilizes polarization multiplexing and filtering, and latching and bidirectional logic. The combination of 2-D arrays of compare-and-exchange modules with optical perfect-shuffle interconnections could lead to pipelined optical sorting networks that can process large numbers of high-bandwidth signals in parallel. Optical sorting networks with these characteristics are applicable in telecommunication switches, parallel processor interconnections, and database machines.

Many materials exhibit a nonlinear change in the index of refraction, which measures how fast an optical wavefront travels through a medium, and which depends on the laser light intensity. Semiconductor materials have received the most attention for application in etalons, because they interact with light in short distances [1.1]. The semiconductors already studied include gallium arsenide (GaAs) and gallium arsenide-aluminum gallium arsenide (GaAs-AlGaAs) multiple quantum wells (MQWs) (Section 3), indium antimonide (InSb), indium arsenide (InAs), copper chloride (CuCl), cadmium sulfide (CdS), and cadmium mercury telluride (CdHgTe). These materials have very large nonlinear indices of refraction, attributable to electronic mechanisms near the edge of their band-to-band absorption. The GaAs-based devices (Section 4) are the most promising for application in optical signal processing because: they operate at room temperature; they are very fast, e.g., NOR-gate response in 1 to 3 ps with < 100-ps recovery time; and they are compatible with diode lasers.

Other materials exhibit thermally-induced changes in both their refractive index and physical thickness. Inside an etalon structure, these effects can combine to produce optical bistability. These materials include semiconductors such as GaAs, zinc sulfide (ZnS), and zinc selenide (ZnSe), as well as color filters and liquid dyes. Of particular interest are the ZnS and ZnSe materials (Section 2). Devices are made from these materials in the form of conventional IFs, in which the semiconductor is the spacer layer. Although much slower than the GaAs etalons, the IFs operate with visible light and are more easily fabricated (if the spacer is polycrystalline). At Heriot-Watt University, three nonlinear interference filters have been combined in a lock-and-clock system, where information is advanced around the ring by an electronic computer [1.14]. At the University of Arizona in Tucson, such filters have been used as decisionmaking devices in one-bit addition by symbolic substitution and in the recognition of particular patterns of three bright beams [1.15, 1.16]

(Section 13); for fingerprint identification using a Van der Lugt correlator and an interference filter as a reflective thresholding device [1.16] (Section 14); and in all-optical compare-and-exchange operations using ZnS bistable optical devices [1.17] (Section 15).

TABLE OF CONTENTS

A.	ZnS Interference Filters and GaAs Etalon Device Research	1
1.	Introduction: Controlling Light with Light	1
1.1	Bistable Devices as Memories and Transistors	1
1.2	Bistable Systems and Their Physics	2
1.3	Two-Wavelength Gate Operation	4
1.4	Nonlinear Materials for Practical Bistable Devices	6
2.	ZnS Interference Filters as All-Optical Logic Gates	8
2.1	Introduction	8
2.2	Theoretical Basis	8
2.3	Design Considerations	13
2.4	Characterization and Results	24
2.5	Discussion	30
3.	Optical Nonlinearity of GaAs and GaAs/AlGaAs Multiple-Quantum Wells	33
3.1	Introduction	33
3.2	Experimental Results	33
3.3	Discussion of Results	35
3.4	Conclusion	40
4.	Uniformity of GaAs Etalons	41
4.1	External Mirrors	41
4.2	Integrated Mirrors	44
4.3	Arrays With Integrated Mirrors	48
4.4	Integrated Mirrors: Nonlinear Optical Performance	49
4.5	Integrated Mirrors: Reduction in Thermal Sensitivity	52
5.	Cascading and Differential Gain of GaAs Etalons Used As Logic Gates	56
5.1	Introduction	56
5.2	Definitions	58
5.3	Experimental Setup	59
5.4	Discussion of Experimental Results	61
5.5	Possible Applications	65
5.6	Conclusion	66
6.	Theoretical Analysis of Differential Gain in GaAs Etalons as a Function of Pulse Length	67
7.	Enhancement of Dynamic Differential Gain of GaAs Etalons by Angle Timing of the Switch Beam	71
7.1	Introduction	71
7.2	Including Angle-Dependence in the Fabry-Perot Formula	72
7.3	Differential Energy Gain as a Function of Internal Angle Between Beams	73
7.4	Effect of Angle-Tuning on Switching Intensity	75
7.5	Assessment of Plane-Wave Approximation and Beam Walk-off	76
7.6	Conclusions	78

8.	Fan-In Of Signals to a GaAs Etalon Device	79
8.1	Introduction	79
8.2	Experiment	80
9.	Direct Fiber-Etalon-Fiber Interfacing	83
9.1	Introduction	83
9.2	Fiber-Etalon-Fiber Interfacing in Transmission	83
9.3	Fiber-Etalon Coupling in Reflection Using Device With MBE-Grown Mirrors	87
9.4	Discussion	89
9.5	Conclusions	90
10.	Holographic Optical Interconnects	91
10.1	Introduction to Optical Interconnects	91
10.2	Binary Phase Grating as CGH	97
10.3	Thick Phase Hologram	105
10.4	Experiments on Planar Thick Holograms	113
10.5	HOE Generating Multiple Equal Focusing Beams for Optical Interconnects	121
11.	Use of a Holographic Optical Element For Fan-Out of Signals to a GaAs Etalon	129
B.	Comparison of Devices for Optical Signal Processing	133
12.	Achievements Versus Statement of Work: Prospects for Nonlinear Decisionmaking in Optical Signal Processing	133
12.1	Introduction	133
12.2	Interference Filters	133
12.3	GaAs Etalons	134
12.4	Application Demonstrations	137
12.5	Comparison of ZnS Interference Filters and GaAs Etalons	137
12.6	Problems With Using Nonlinear Optical Devices	138
C.	Optical Signal Processing Using ZnS Interference Filters for Nonlinear Decisionmaking	142
13.	Pattern Recognition and Symbolic Substitution	142
13.1	Introduction	142
13.2	Component Design and IF Performance	143
13.3	Pattern Recognition	150
13.4	One-Bit Adder Experiment	156
13.5	Discussion	159
14.	Associative Memory of Fingerprint Identification	161
14.1	Motivation and Concept	161
14.2	Experimental Apparatus	164
14.3	Results	168
14.4	Some Practical Details	173
14.5	Summary and Discussion	177

15.	All-Optical Compare-and-Exchange Switches	179
15.1	Introduction	179
15.2	Compare-and-Exchange Circuit Design	181
15.3	Experimental Demonstration	183
15.4	Discussion	188
	References	194
APPENDIX A:	Room-Temperature Excitonic Optical Nonlinearities of Molecular Beam Epitaxially Grown ZnSe Thin Films	204
APPENDIX B:	Introduction to Dichromated Gelatin (DCG) Recording Material	212
APPENDIX C:	Conclusions from the 1988 PhD dissertation of George Gigioli, Jr., entitled "Optimization and Tolerancing of Nonlinear Fabry-Perot Etalons for Optical Computing	218

A. ZnS INTERFERENCE FILTERS AND GaAs ETALON DEVICE RESEARCH

1. INTRODUCTION: CONTROLLING LIGHT WITH LIGHT

1.1 Bistable Devices As Memories and Transistors

A system is said to be optically bistable if there are two stable output intensities possible for the same input intensity [1.1]. Thus, in Fig. 1.1a, the output intensity, I_T , can have two values for any input holding intensity, I_H , lying above the switch-down intensity, I_L , and below the switch-up intensity, I_U . Such a bistable system can be used as an optical memory element, but a continuous input beam of intensity I_H must be maintained. One can observe the transmitted intensity, I_T , as the input intensity, I_I , is slowly increased from zero to well above the switch-up intensity, and then decreased below the switch-down intensity; this maps out the bistable hysteresis curve of I_T versus I_I , as shown schematically in Fig. 1.1a and for GaAs in Fig. 1.2 [1.3]. Such optical bistability in a passive nonlinear Fabry-Perot interferometer was first observed in 1974 by H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan at Bell Labs using sodium vapor inside a 11-cm-long interferometer [1.4].

The bistability hysteresis curve can be made very narrow or made to barely disappear, yielding a region of very steep dependence of I_T upon I_I (see Fig. 1.1b). In the latter case, the device operates as an all-optical transistor [1.4]. Over some range of amplitude modulation of I_I about its value I_g for maximum gain, the device linearly amplifies the modulation; i.e., a weak beam controls a more intense beam. For a small change δ in I_I , the device exhibits thresholding: for I_I below $I_g - \delta$ the transmission is quite low, and for I_I above $I_g + \delta$ the transmission is high. A threshold device can serve as an optical discriminator, optical limiter, or optical AND gate. A bistable device can be operated as an AND gate between a holding beam I_H less than $I_g - \delta$ and a switching beam such that the sum exceeds $I_g + \delta$. Often the switching beam is a pulse; if so, its duration must exceed the cavity and nonlinear medium relaxation times if the steady-state characteristics shown in Figure 1.1 are to be valid. The holding beam maintains the device in the lower state—the switch is "off." When the switching beam is added, the transmitted intensity jumps to the higher state, and the switch turns "on." In the "off" state, very little light is transmitted through the device; in the "on" state, light transmission is high. If there are two input pulses, I_H can be set for either AND-gate or OR-gate operation.

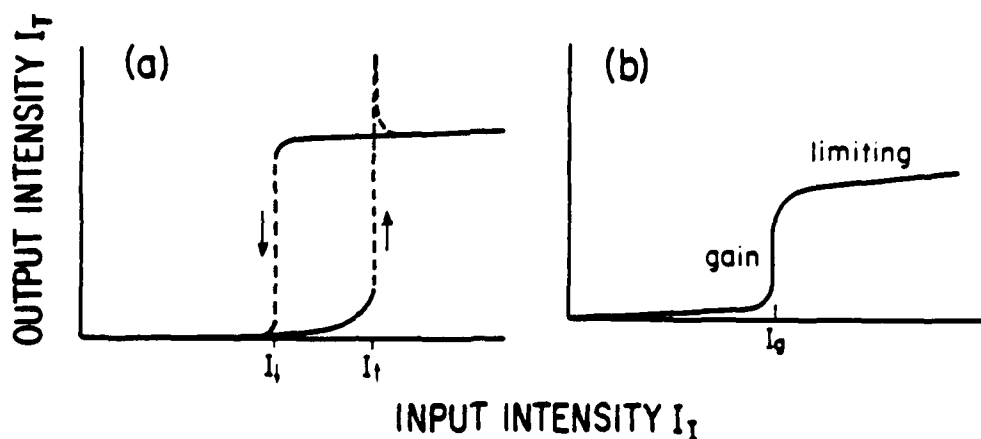


Figure 1.1 The output-input characteristic of a nonlinear etalon. In (a) optical bistability is shown where there are stable transmission states (high and low transmissions) between I_l and I_h . In (b), thresholding occurs.

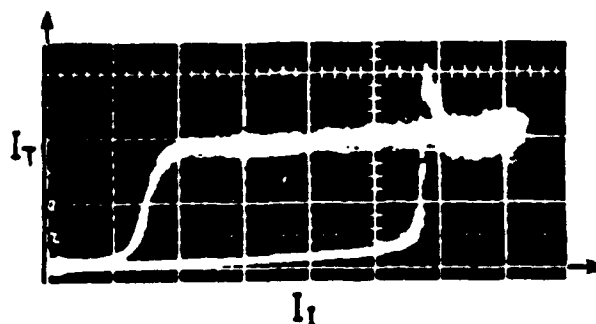


Figure 1.2 Optical bistability at room temperature in a GaAs/AlGaAs MQW etalon.

1.2 Bistable Systems and Their Physics

Consider the examples of bistable systems shown in Fig. 1.3: a nonlinear Fabry-Perot interferometer, a nonlinear ring cavity, and a hybrid bistable device. In general, a Fabry-Perot interferometer consists of two mirrors aligned to maximize the interference of multiple reflections between the two mirrors. The most promising optical bistable devices, at present, are very thin, solid Fabry-Perot interferometers called etalons. The optical pathlength of any sample is the physical length multiplied by the refractive index of the material. If the optical pathlength of the etalon is equal to an integral number of half wavelengths of the incident light, the etalon is in resonance and highly transmits the incident light. But in a *nonlinear* etalon, the resonance peak shifts as the light intensity changes. If the feedback into the

nonlinear material from the partially reflecting mirrors is high enough, the etalon can exhibit optical bistability. This operation is achieved most easily when the laser is detuned at low intensities by one resonance width from the Fabry-Perot resonance peak. At sufficient intensities, the nonlinear optical pathlength shifts the peak into resonance and slightly beyond--in a positive-feedback runaway manner. The intensity inside an on-resonance etalon is higher than that outside because of its storage property. Therefore, it is not surprising that once the etalon is in resonance, it can be maintained in resonance with less input than was required to shift it into resonance; this is the origin of the bistability hysteresis.

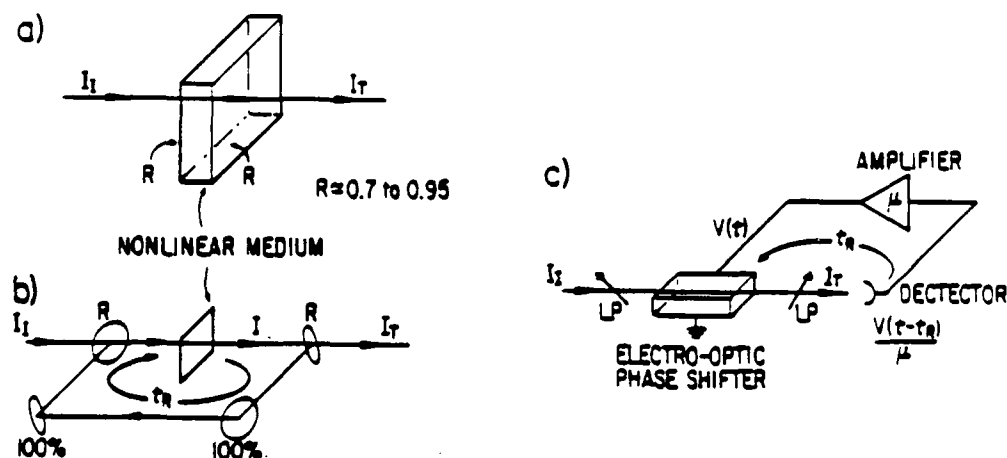


Figure 1.3 (a) Nonlinear Fabry-Perot etalon consisting of solid material with parallel surfaces with coatings of reflectivity R . (b) Nonlinear ring cavity consisting of three or four mirrors so the light propagates in only one direction in the nonlinear medium. (c) Hybrid bistable device in which the nonlinearity involves conversion of the transmitted intensity to an electrical signal, which is amplified and fed back to the electro-optic crystal, changing the light transmitted through the crossed polarizers LP . The feedback delay, τ_R , is kept to a minimum in bistability experiments, but it is made longer than the system response time for delayed-feedback optical chaos experiments.

In a Fabry-Perot interferometer, the forward and backward fields interfere to form a standing wave with a spatial period of half a wavelength of the light in the material. Sometimes free motion or diffusion averages out the effects of this rapid variation on the polarization, so that it can be neglected. However, theorists prefer to eliminate the standing wave altogether, by using a nonlinear ring cavity (refer to

Fig. 1.3b), in which the light propagates through the medium in only one direction [1.5]. Kimble and co-workers at the University of Texas saw bistability in such a system in 1983 [1.6]. It is the physical system most often studied numerically.

Peter Smith and E. H. Turner at Bell Labs realized in 1977 that the nonlinear material in a nonlinear Fabry-Perot interferometer could be replaced by an artificial nonlinearity [1.7]. The intra-cavity index was made intensity dependent by detecting I_T , amplifying the detector's output voltage, and applying the resulting intensity-dependent high voltage to an intra-cavity electro-optic crystal. Garmire, et al. realized that a cavity is not essential for bistability and demonstrated hybrid bistability using an electro-optic crystal between crossed polarizers (see Fig. 1.3c) [1.8]. This device has been particularly useful in fundamental studies of critical slowing and instabilities because it is truly a plane-wave device. That is, the phase shift is the same across the entire Gaussian transverse profile of the laser beam. In an all-optical device, the pathlength change is determined by the local light intensity (unless rapid diffusion or conduction averages out this effect). David Miller, et al. are developing self electro-optic effect devices (SEEDs), based on the increase of optical absorption of a GaAs-AlGaAs MQW structure with application of an electric field perpendicular to its surface [1.9]. SEEDs are slower but require less power than all-optical devices. Arrays of SEEDs are being developed as spatial light modulators and for hybrid optical computing [1.10].

1.3 Two-Wavelength Gate Operation

The AND-gate and OR-gate operations of the nonlinear etalon described above use inputs at a single wavelength. All of the gate operations can be performed using input pulses at one highly absorbing wavelength, to maximize the shift in Fabry-Perot peak for a probe (output) pulse at a second wavelength (see Fig. 1.4) [1.11]. Various logic gates such as AND, OR, and NOR have been demonstrated in GaAs etalons, as exhibited in Fig. 1.5 [1.12]. A GaAs optical-logic NOR gate has been operated at a 100-MHz rate with a few picojoules switching energy per pulse. The answer to the NOR-gate question is obtained in a few picoseconds and can be on its way to the next gate long before the first gate is ready to be used again. A recovery time as short as 50 ps was demonstrated (corresponding to 5 billion operations per second) in 1986 by Jewell, Lee, et al. at Bell Labs, but heating effects may limit the rate to 1 billion operations per second or less [1.13]. This type of logic-gate operation is the most efficient, but it has the disadvantage that the input and output signals have different wavelengths, complicating the design of a system with sequential logic

operations. Less efficient single-wavelength bistable or transistor operation of a nonlinear etalon is possible, as described earlier in this section.

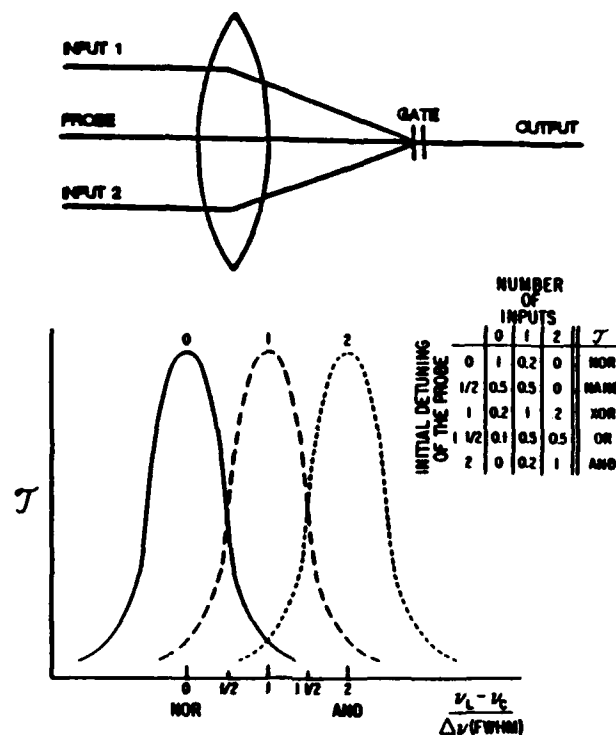


Figure 1.4 Use of a nonlinear etalon to perform AND and NOR logic operations. The two logic inputs are at a wavelength that is absorbed well by the medium, so they are very efficient in changing the refractive index and shifting the Fabry-Perot peak transmission. The curve labeled 0 is the probe's transmission as a function of frequency if no input pulses are present. Its transmission is high at a frequency labeled 0 and falls to one-half of its peak value at a frequency labeled 1/2. If only one input is present, the probe's transmission shifts to the dashed curve labeled 1. If both inputs are present it shifts to 2. The probe pulse follows just after the inputs. Its transmission can be high or low, depending upon its frequency and the number of inputs that precede it. If its transmission is high, the answer is 1, or TRUE; inversely, if its transmission is low, the logic output is 0, or FALSE. For example, to achieve an AND gate, the probe pulse has frequency labeled "AND." If either or both inputs are missing, the answer is FALSE (low probe transmission). But if both inputs are present, the answer is TRUE (high probe transmission). For a NOR gate, the probe frequency is set at "NOR." Its transmission is high with no inputs, but low with either or both inputs. All the gates can be achieved with appropriate probe frequency; see inset table.

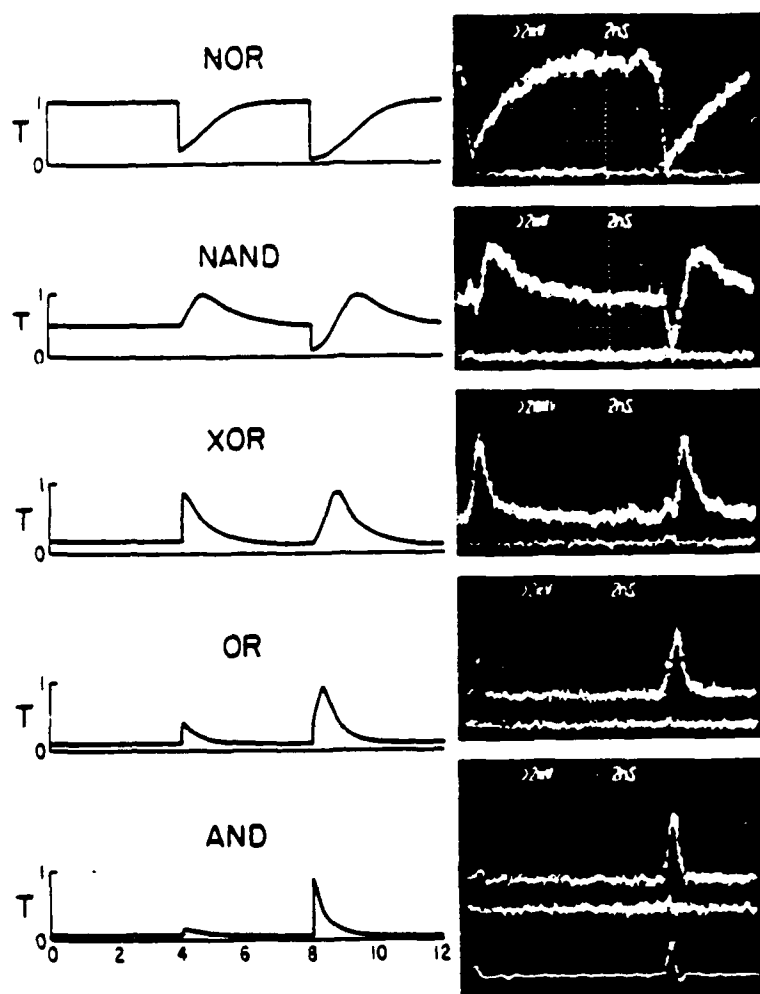


Figure 1.5 Optical logic gate operation. Left: computer simulations of the transmission \mathcal{T} of a cw probe as a function of time, with one input occurring at time 4 and two inputs occurring at time 8. Use of a cw probe permits observation of the recovery of the gate as well. Right: cw probe transmission showing gate operations using a GaAs/AlGaAs MQW etalon. The straight-line trace in each photograph shows the zero-transmission level of the etalon (probe blocked). The inputs (at $\cong 0$ ns and $\cong 14$ ns) are shown as the bottom trace in the AND photograph.

1.4 Nonlinear Materials for Practical Bistable Devices

Although the basic principle of an optical transistor has been understood for a couple of decades, it was not until the late 1970s that huge optical nonlinearities were discovered in indium antimonide (InSb) at 77 K by S. D. Smith, D. A. B. Miller, et al. at Heriot-Watt University in Edinburgh, and in GaAs between 4 K and 120 K by H. M. Gibbs, S. L. McCall, A. C. Gossard, et al. at Bell Labs [1.1]. It was 1982 before low-power, room-temperature bistability was seen in GaAs (see

Figure 1.2). Researchers continue to search for materials with large nonlinearities and low linear losses at room temperature.

Many materials exhibit a nonlinear change in refractive index; i.e., the index of refraction, which measures how fast an optical wavefront travels through a medium, and which depends on the laser light intensity. Semiconductor materials have received the most attention for etalon applications because they interact with light in short distances [1.1]. The semiconductors already studied include GaAs, and gallium arsenide-aluminum gallium arsenide (GaAs-AlGaAs) MQWs (Section 3), InSb, indium arsenide (InAs), copper chloride (CuCl), cadmium sulfide (CdS), and cadmium mercury telluride (CdHgTe). These materials have very large nonlinear indices of refraction attributable to electronic mechanisms near the edge of their band-to-band absorption. The GaAs-based devices (Section 4) are the most promising for applications: they operate at room temperature; they are very fast, e.g., NOR-gate response in 1 to 3 ps with < 100-ps recovery time; and they are compatible with diode lasers.

Other materials exhibit thermally induced changes in both their refractive index and physical thickness. Inside an etalon structure, the effects can combine to produce optical bistability. These materials include semiconductors such as GaAs, zinc sulfide (ZnS), and zinc selenide (ZnSe), as well as color filters and liquid dyes. Of particular interest are the ZnS and ZnSe materials (Section 2). Devices are made from these materials in the form of conventional interference filters, in which the semiconductor is the spacer layer of the filters. Although much slower than the GaAs etalons, the interference filters operate with visible light and are more easily fabricated (if the spacer is polycrystalline). At Heriot-Watt University, three nonlinear interference filters have been combined in a lock-and-clock system, where information is advanced around the ring by an electronic computer [1.14]. At the University of Arizona in Tucson, such filters have been used as the decisionmaking devices in one-bit addition by symbolic substitution and in the recognition of particular patterns of three bright beams (Section 13), [1.15, 1.16]; for fingerprint identification using a Van der Lugt correlator and an interference filter as a reflective thresholding device (Section 14), [1.16]; and all-optical compare-and-exchange operations using ZnS bistable optical devices (Section 15), [1.17].

2. ZnS INTERFERENCE FILTERS AS ALL-OPTICAL LOGIC GATES [2.1]

2.1 Introduction

The use of optical bistable devices in optical computing is envisioned in direct analogy with the use of transistors in electronic counterparts. Many III-V and II-VI semiconductors have been investigated as potential candidates for an optical transistor since large optical nonlinearities in the vicinity of band edges can be achieved. In a ZnS etalon, laser-induced heating shifts the absorption band edge to the longer wavelength. By Kramers-Kronig technique, one can show that this produces refractive index changes large enough to shift the transmission peak of the etalon; therefore, optical bistability is obtained as described in Section 1. Although this $\partial n/\partial T$ thermal process gives ZnS devices switch on/off times typically in the 10 μ s-1 ms range, which is quite slow compared with GaAs devices, the fabrication of ZnS devices involves common, easy-to-access evaporation equipment rather than relatively expensive molecular beam epitaxy (MBE) facilities. This is particularly convenient in making devices for optical computing demonstrations. In this section, dispersive optical bistability and the graphical method of its solutions are briefly reviewed. Linear and nonlinear considerations in the design of ZnS interference filters (IFs), in terms of peak transmission, full-width-half-maximum (FWHM), switching power, detuning, etc. are presented. Finally, experimental data obtained from the linear and nonlinear measurements are compared with the design work.

2.2 Theoretical Basis

Assume a Fabry-Perot cavity is filled with a nonlinear material and the input and output mirrors have the identical intensity reflectance R . The two mirrors are assumed to be *linear and lossless*, so that $R = 1 - \mathcal{T}_m$. After the n th round trip, the electric field just inside the input mirror of the Fabry-Perot cavity ($z = 0$ in Fig. 2.1) by mean-field theory can be written as

$$E_{n+1}(0) = E_n(0)e^{(ik-\sigma)2l}R + E_I\sqrt{\mathcal{T}_m}. \quad (2.1)$$

For steady state, let $E_{n+1}(0) = E_n(0) = E_{sp}$; then we have

$$E_{sp} = \frac{E_I \sqrt{\mathcal{F}_m}}{1 - \text{Re}(ik - \sigma)2l} \quad (2.2)$$

and

$$\frac{E_T}{E_I} = \frac{\sqrt{\mathcal{F}_m} E_{sp} e^{(ik - \sigma)l}}{E_I} = \frac{\mathcal{F}_m e^{(ik - \sigma)l}}{1 - \text{Re}(ik - \sigma)2l} \quad (2.3)$$

where l is the length of the cavity, k is the propagation constant and $\sigma = \sigma_r + i\sigma_i$ is the complex-field absorption coefficient. Here the conventional intensity absorption coefficient α is related to σ by $\alpha = 2\sigma_r$. For two-level atoms, both the real part σ_r and the imaginary part σ_i depend on the applied intensity and detuning [2.2]. For a complete review of the models of optical bistability, see Gibbs [2.3].

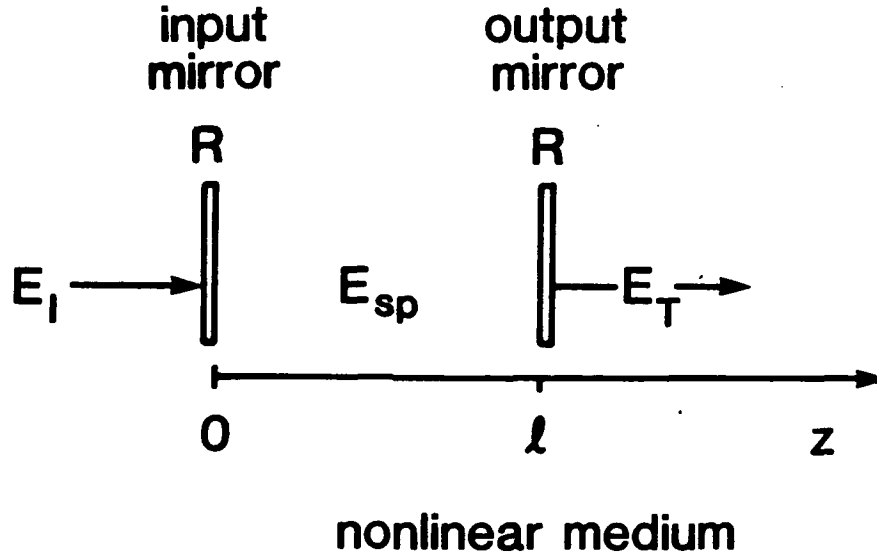


Figure 2.1 Lossless nonlinear Fabry-Perot cavity. E_I : input field; E_{sp} : mean cavity field; E_T : transmitted field.

In general, absorptive optical bistability is difficult to observe and to use, while dispersive optical bistability has been observed in many materials and gives greater promise for application. Therefore, discussion here is limited to the dispersive case.

For example, in a Kerr medium or a system of two-level atoms with detuning far from resonance, one effectively has purely dispersive bistability in which $\sigma_r = 0$. From Eq. (2.3),

$$\frac{I_T}{I_I} = \frac{|E_T|^2}{|E_I|^2} = \frac{\mathcal{F}_m^2}{|1 - R e^{i\beta}|^2} = \frac{1}{1 + \frac{4R}{\mathcal{F}_m^2} \sin^2 \left[\frac{\beta}{2} \right]}. \quad (2.4)$$

$$\begin{aligned} \beta &= (k - \sigma_i)2l \\ &= \beta_0 + \beta_2 I_T \end{aligned} \quad (\text{in Kerr medium}), \quad (2.5)$$

where β is the round-trip phase shift, and β_0 contains all non-intensity-dependent phase shifts, including initial detuning and the phase change at interfaces as light propagates through the Fabry-Perot cavity. From Eq. (2.5), the transmission can also be expressed as

$$\frac{I_T}{I_I} = \frac{\beta - \beta_0}{\beta_2 I_I}. \quad (2.6)$$

A plot of I_T/I_I versus I_T with certain values of β is given in Fig. 2.2 and the corresponding plot of I_T versus I_I is shown in Fig. 2.3. The graphical approach, giving possible solutions of the transmitted intensity for a given input intensity, is described briefly as follows [2.4]. The straight lines in Fig. 2.2 represent various values of I_I , and as indicated by Eq. (2.6), the steeper a line's slope the lower the input intensity. The intersection points between these straight lines and the Airy-function curve are used to generate the I_T versus I_I plot in Fig. 2.3. Bistability occurs where the straight line has three intersections. The upper branch of a bistable loop in Fig. 2.3 corresponds to the negative-slope side of the Airy-function curve in Fig. 2.2, while the unstable state shown in dashed line in Fig. 2.3 corresponds to the positive-slope side of the Airy-function curve. The instability means that a small change in the input field E_I will cause the output to evolve to either the upper or the lower stable branch.

If $|\beta| \ll 1$, e.g., in the vicinity of a transmission peak, Eq. (2.4) then gives

$$\frac{I_T}{I_I} \approx \frac{1}{1 + \frac{R}{\mathcal{F}_m^2} (\beta_0 + \beta_2 I_T)^2}. \quad (2.7)$$

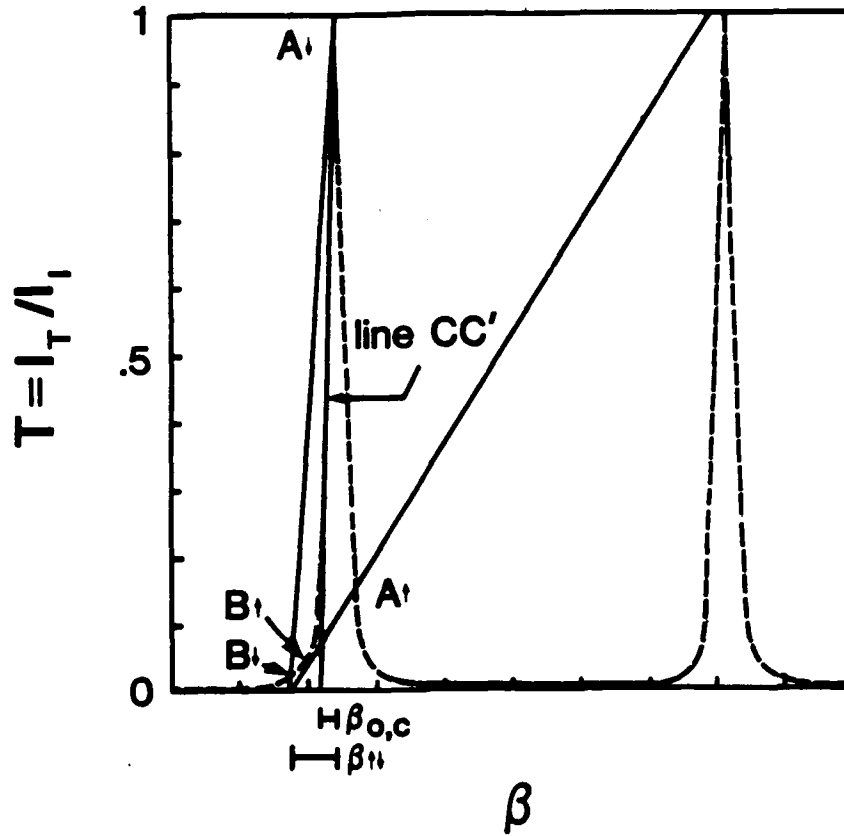


Figure 2.2 The curve is a plot of the Fabry-Perot transmission, Eq. (2.4), vs. the phase shift β . The straight lines represent different input intensities I_I (see text). Line A_1B_1 for switch-up and A_2B_2 for switch-down; line CC' for critical switching.

For bistability to occur, one requires that $\frac{dI_I}{dI_T} > 0$. At the onset of bistability and at the inflection point, $\frac{dI_I}{dI_T} = 0$ and $\frac{d^2I_I}{dI_T^2} = 0$, respectively. Thus, by Eq. (2.7) one finds a quadratic equation of $\beta_2 I_T$ which has two roots:

$$\beta_2 I_T = -\frac{2}{3\beta_0} \pm \frac{1}{3 \sqrt{\beta_0^2 - \frac{3\mathcal{G}_m^2}{R}}} \quad \text{for } \frac{dI_I}{dI_T} = 0. \quad (2.8a)$$

$$\beta_2 I_T = (-2/3)\beta_0 \quad \text{for } \frac{d^2I_I}{dI_T^2} = 0. \quad (2.8b)$$

Consequently, conditions of the occurrence of bistability are derived by requiring that the argument in the square root be real (2.8a) and an intensity be positive (2.8b). We have

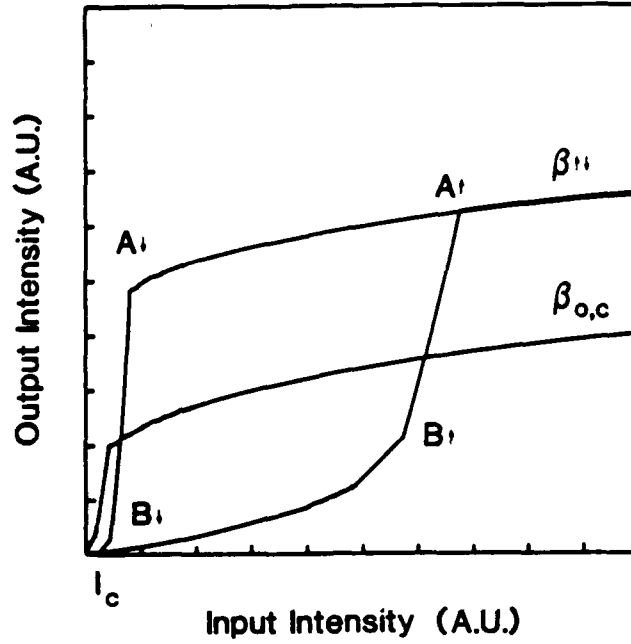


Figure 2.3 Bistable loops. I_c : critical input intensity for the onset of bistability.

$$\begin{aligned} \text{(i)} \quad |\beta_0| > \sqrt{\frac{3}{R}} \mathcal{F}_m = \frac{\sqrt{3}\pi}{\mathcal{F}} = \frac{\sqrt{3}}{2} \Delta_{1/2, \text{phase}} \cong \Delta_{1/2, \text{phase}} \\ \text{(ii)} \quad \text{sign}(\beta_0) = - \text{sign}(\beta_2) \end{aligned} \quad (2.9)$$

where $\mathcal{F} = \frac{\pi\sqrt{R}}{1-R}$ is the cavity finesse and $\Delta_{1/2, \text{phase}} \equiv \text{FWHM}_{\text{phase}} = 2\pi/\mathcal{F}$. The first condition indicates that the initial detuning should be larger than one FWHM of the Fabry-Perot cavity to obtain optical bistability. A numerical example is given: if $R = 0.9$, then $\mathcal{F} \cong 30$ and $\Delta_{1/2, \text{phase}} = 0.21$. Referring to Eq. (2.9) and Eq. (2.8b), the higher the finesse, the smaller the initial detuning is required; the smaller the nonlinearity β_2 , the larger I_1 must be.

At a certain detuning β_{11} , corresponding to the maximum input power that can be applied to the etalon, the switch-up or switch-down intensity is determined from the intersections of the Airy-function curve and line A_1B_1 or A_1B_1 , respectively. The difference between two slopes of these two lines determines the range of input intensity, i.e., the width of a bistable loop, for which a nonlinear etalon has two stable outputs. One criterion in determining the performance of a nonlinear etalon is the critical input intensity for the onset of bistability I_c . It is defined as the

minimum incident intensity below which bistability cannot be observed, regardless of the initial detuning β_0 [2.5]. This is shown by a line CC' in Fig. 2.2, which has the slope equal to the maximum gradient of the Airy-function curve. The line CC' meets the x-axis at a point which is separated from the point corresponding to the peak transmission by an amount $\beta_{0,c}$, referred to as the critical initial detuning. More discussions about the critical input intensity and the critical detuning will be in the next subsection. Multistability can occur if line A_1B_1 intersects more than one transmission peak.

Resonance enhancement in a narrow-band IF provides the electric field necessary to observe nonlinear effects of optical materials. In GaAs or InSb, this enhancement causes an electronic nonlinearity which leads to fast and low-power-consumption devices, see Section 4. In ZnS or ZnSe, the enhancement causes a thermal nonlinearity. Karpushko, et al. [2.6-2.8] were the first observers of optical bistability in thin-film etalons and attributed the mechanism to two-photon photorefraction. Later, Olbright, et al. [2.9] indicated that a one-photon thermal mechanism is responsible for the optical bistability in ZnS or ZnSe IFs, evidenced by: the lack of strong excitonic peaks in the spectrum of ZnS material at room temperature; the positive sign of nonlinear index change; and the time scale of thermal switching. The *absorption* properties of the ZnS or ZnSe cavity cause heating which shifts the absorption edge to longer wavelength and the induced temperature increase ΔT is proportional to the input intensity. Heating has the effect of changing the optical path length of a cavity.

$$\frac{\Delta OPL}{\Delta T} = \frac{\Delta(nl)}{\Delta T} = \frac{l\Delta n}{\Delta T} + \frac{n\Delta l}{\Delta T}, \quad (2.10)$$

where $n = n_0 + n_2 I$. Because $l\Delta n/\Delta T \gg \Delta l/\Delta T$ ($l\Delta n/\Delta T \cong 10^{-4} \mu m/C^\circ$ and $n\Delta l/\Delta T \cong 10^{-6} \mu m/C^\circ$ for bulk ZnS at $\lambda = 514.5 \text{ nm}$), $\Delta OPL/\Delta T \cong l\Delta n = (n_0 + n_2 I)/\Delta T$. This intensity-dependent phase change can tune the transmission peak of a nonlinear etalon from off-resonance onto resonance, switching the etalon from low to high transmission.

2.3 Design Considerations

In general, a narrow-band IF consists of a spacer layer surrounded by quarter-wave stacks as high-reflection mirrors, all deposited on a substrate. This structure can be represented by a formula: $G|(HL)^p m(HH) (LH)^p |g|G$, where H and L represent a quarter-wave layer of High/Low refractive index material respectively, m

is the order of the spacer, p is the order of the high-reflectance mirrors, g is an optical glue and G is the glass for cover and substrate. In our experiments, ZnS or ZnSe is chosen for H, cryolite (Na_3AlF_6) for L, and a $50 \times 50 \times 1$ mm³ Kodak Slide cover glass ($n = 1.52$) for G. The total physical thickness of the multilayer structure typically is $\approx 1\text{-}2$ μm .

The technique of thermal vapor deposition has been employed to make all of the IFs in the Optical Coating Laboratory of the Optical Sciences Center using an Edwards box coater at a vacuum of $\approx 6 \times 10^{-6}$ Torr. Optical glues are applied to protect the IFs. The reasons are as follows. A deposited film can be observed by electron microscopy to show the columnar structure, instead of the solid planar structure [2.10]. Voids between columns cause an inhomogeneous broadening and an exponential extension (Urbach tail) of the band edge toward lower energies, giving absorption in the band-tail spectral region considerably higher than that in the corresponding bulk materials. Therefore, absorption of moisture ($n \approx 1.3$) into the previously evacuated voids causes a drift of an IF's peak wavelength towards longer wavelengths as the refractive index increases. Moreover, chemical reactions of the adsorbed water with the film's material are suspected to further degrade the performance of an IF to the extent that sometimes the nonlinearity can not even be observed for an unprotected IF at the IF's design wavelength. Therefore, to prevent moisture penetration and drift, Norland 61 UV-curing cement is used for sealing purposes.

There are two aspects regarding thin-film properties that need to be mentioned before discussing design considerations of the nonlinear IFs.

(i) Thin films have different characteristics from the bulk counterparts such as a reduced refractive index, lower thermal conductivity, broader band edge, which make it difficult to accurately know the value of a refractive index and an absorption coefficient of the material in a film. So the well-tabulated data of bulk materials can only be used as informative references.

(ii) Different manufacturing conditions, e.g., residuals in the vacuum or an impurity in the material to be deposited, can alter the residual absorption and scattering at film interfaces, which may affect an IF's performance. Therefore, the finesse of an etalon is not necessarily improved as the reflectance of mirrors increases because of the additional losses.

Although thin-film structures do contribute these complications, the results obtained from the design work will be instructive in the sense that they can be used

reasonably well as guidelines for the experimental work. However, sometimes quantitative predictions might not exactly match the experimental data and further discussion is needed.

The design considerations are based on the requirements of optical switches such as power consumption, fan-out, contrast, speed, etc. For ZnS IFs, the design includes both the linear and the nonlinear parts. Uniformity of peak transmission, peak wavelength and FWHM over a large area on a ZnS IF can be regarded as the design goal of the linear part while low switching power and fast switching time are the goals of the nonlinear part. There are certainly some tradeoffs among design considerations of both the linear and the nonlinear part, which will be discussed in the following sections. Additionally, a filter requires perpendicular operation, particularly in massive parallel computation, to prevent different spot sizes on a ZnS IF under different focusing conditions. This requires that the peak wavelength of a ZnS IF be shorter than the wavelength of a source, as can be realized by using an oblique monitoring beam of the same laser in the deposition process.

One important linear device characteristic for parallel processing is uniformity, which depends on the coating geometry and the monitoring technique. Macleod shows that the thickness of deposit on the substrate decreases monotonically with the distance away from the line connecting a point source and the center of a flat plane substrate [2.11]. This indicates that the peak wavelength measured near the edge of an IF will be inherently shorter than that at the center. To improve the uniformity, the method of rotating a substrate during deposition has been proven successful because different positions on the substrate will see an equal average growth rate during the deposition. Thickness control in the process of thin-film deposition is vital to the performance of an IF. An optical monitoring technique is used to detect the transmittance of the depositing film, which generally exhibits a turning point for each quarter-wave optical thickness. This technique has the advantage that the calibration of an optical system is not necessary as it is in the quartz-crystal method; however, the optical system requires a same-wavelength light source. In addition, an operator must be experienced at determining when the exact turning point is reached and know when to terminate the layer.

Variations in the thickness of thin-film layers, often encountered in real deposition, will affect the peak wavelength of a ZnS IF. For example, a thicker spacer will shift the peak wavelength to the longer side and vice versa because of the $2n/\cos\theta = m\lambda$ relationship. If one bears this in mind, a thickness error in a layer, e.g., overshoot, can be compensated for by adjusting the thickness of the next layer. The sensitivity of the peak wavelength to variations in the thickness of a

layer can be simulated by a program developed by Macleod. It is concluded that, first, deviations in thickness occurring at mirror layers close to the spacer will have a greater effect than those farther from the spacer; second, a deviation in the spacer will have a greater effect than one in the mirrors. In addition, the program can also include absorption in each layer. Calculations of different peak transmissions and FWHMs due to different combinations of spacer thickness and mirror reflectance have been performed, and the results are summarized in Table 2.1. The ZnS IFs are designed using previously published values for necessary constants, e.g., at $\lambda = 514.5$ nm, $n \cong 2.419$, $\alpha \cong 130$ cm⁻¹ for ZnS, $n \cong 1.35$, $\alpha = 0$ cm⁻¹ for cryolite [2.9, 2.12]. In the calculation, two cases are compared when the absorption is taken into account: absorption is included in all H layers, and absorption is included only in the spacer layers. The former case is calculated by the program mentioned above; the latter case can be expressed as follows for peak transmission and FWHM.

$$\mathcal{T}_{\text{peak},s} = \frac{(1-R)^2 e^{-\alpha l}}{(1-Re^{-\alpha l})^2} \quad (2.11)$$

$$\text{FWHM}_s = \frac{\lambda_0}{(m+p)\pi} \frac{1-Re^{-\alpha l}}{\sqrt{Re^{-\alpha l}}} \quad (2.12)$$

where m is the order of the Fabry-Perot transmission curve and is equal to the order of the spacer, p is the phase change factor $\cong 1$ -1.5 with ZnS and cryolite as the material [2.13], and R is the reflectance of a mirror stack which will be discussed later. The conclusions are that (i) the spacer thickness of a ZnS IF determines its peak wavelength and has a weak influence on the FWHM, as expected; (ii) the reflectance of the dielectric mirrors strongly affects the peak transmittance and FWHM; (iii) absorption in all layers *does* change the peak transmission and FWHM, as indicated in Table 2.1.

Observation of optical bistability often requires a focused light beam to obtain enough intensity. The effect of the incident cone of light on the Fabry-Perot can be summarized as follows [2.14].

$$\Delta\lambda' \cong \lambda_\theta - \lambda_0 \cong \frac{\lambda_{0,\text{peak}}}{8 \cdot (n_H^2 \cdot F/\#)^2} \quad (2.13)$$

$$(\Delta\lambda_\theta)^2 = (\Delta\lambda_0)^2 + (\Delta\lambda')^2 \quad (2.14)$$

and
$$\mathcal{T}_\theta = \mathcal{T}_0 \frac{\Delta\lambda_0}{\Delta\lambda'} \tan^{-1} \frac{\Delta\lambda'}{\Delta\lambda_0} \quad (2.15)$$

Table 2.1 Calculated peak transmissions and FWHMs of ZnS IFs versus the number of spacer layers and mirror layers.*

			p = 3	p = 4	p = 5
(HL) ^p 2H(LH) ^p	T _{peak}	A	90.03%	72.71%	41.40%
		S	98.49%	95.27%	85.96%
	FWHM	A	60 Å	21 Å	8 Å
		S	121-152 Å	38-48 Å	13-16 Å
(HL) ^p 4H(LH) ^p	T _{peak}	A	84.03%	59.94%	26.68%
		S	97.01%	90.86%	74.67%
	FWHM	A	42 Å	15 Å	7 Å
		S	87-102 Å	28-33 Å	10-11 Å
(HL) ^p 6H(LH) ^p	T _{peak}	A	78.60%	50.26%	13.72%
		S	95.56%	86.75%	65.48%
	FWHM	A	34 Å	12 Å	5 Å
		S	69-77 Å	22-25 Å	8-9 Å
(HL) ^p 8H(LH) ^p	T _{peak}	A	73.68%	42.75%	XXXXXX
		S	94.14%	82.92%	XXXXXX
	FWHM	A	30 Å	10 Å	XXXXXX
		S	56-62 Å	19-21 Å	XXXXXX
(HL) ^p 12H(LH) ^p	T _{peak}	A	69.21%	XXXXXX	XXXXXX
		S	92.76%	XXXXXX	XXXXXX
	FWHM	A	24 Å	XXXXXX	XXXXXX
		S	48-52 Å	XXXXXX	XXXXXX

*A: Linear absorption is included in all H layers.

S: Linear absorption is included only in the spacer layers (see discussion in text).

XXXXXX: Impractical region.

where $n_H^* \equiv \sqrt{n_H n_L}$ is the effective index of a high-index spacer material; $\Delta\lambda_0$, $\Delta\lambda_\theta$ are FWHMs of the filter at normal incidence and at an angle θ with respect to normal; \mathcal{T}_0 , \mathcal{T}_θ are peak transmission at normal and at θ . A numerical example is given: assume $F/\# = 5$, $\Delta\lambda_0 \approx 1.5$ nm; then we have $n_H^* \approx 1.807$, $\Delta\lambda_\theta = 1.69$ nm, and $\mathcal{T}_\theta = 0.921\mathcal{T}_0$ at $\lambda_{0,\text{peak}} = 514.5$ nm. Note that a tighter focusing of incident light will result in a broader FWHM, a reduced transmission, and a shift of peak wavelength towards longer wavelength, which in turn will require more detuning in the nonlinear operation of a ZnS IF.

The optimization of switching characteristics for a bistable IF has been studied by several researchers. See D. A. B. Miller [2.5] for an insightful foundation for

discussion. Wherrett [2.15] for a discussion of reflection bistable devices, and Wherrett, Hutchings, and Russell [2.16] for a detailed exploration of the topic. Most of their research was done on ZnSe material because it has a large absorption coefficient, and directed at reflection bistable devices because the analysis shows that the switching power is smaller for any αl if the output mirror becomes more and more reflective. Researchers, however, tend to be more familiar with the use of transmission light than with a reflection beam for cascading, because of the corresponding ease in operation. In addition, cascading with a reflection beam is inconvenient, because the reflection light intensity is low when the device is switched on. The nonlinear design considerations discussed hereafter are based on the theoretical structures mentioned above, but will concentrate on ZnS IFs with reasonably high transmission and low power. ZnSe IFs are mentioned only for the purpose of comparison.

As the number of (HL) stacks for high-reflectance mirrors increases, the mirrors can be assumed to be lossless. For example, mirrors of stacks $H(HL)^2$ and stacks $H(HL)^5$ have $R = 0.8151$ and $R = 0.9927$, respectively, if no loss is included in the calculation. If the absorption is taken into account ($\alpha l \approx 3.2 \times 10^{-3}$ for a single H-layer), $R = 0.7816$ and $R = 0.9901$. As will be seen later, these values suggest that a consideration of the nonlinear material H in the mirror stacks affects the switching intensity only marginally. Therefore, under the assumption of no loss, the reflectance R of a mirror having structure $H(HL)^p | G_{\text{substrate}}$ can be expressed as

$$R = \frac{|(n_H^{2p} - n_s \cdot n_L^{2p})|^2}{|(n_H^{2p} + n_s \cdot n_L^{2p})|^2} \quad (2.16)$$

Further assumptions are: (i) the system exhibits plane-wave operation; (ii) both mirrors have identical reflectance; (iii) linear absorption is considered only in the spacer layers; (iv) the beam spot diameter ($\approx 20 \mu\text{m}$) is greater than the spacer thickness ($\approx 2-3 \lambda$); and (v) for simplicity, temperature is uniform throughout the multilayers.

Referring to Eq. (2.10), the nonlinear phase change of thermal origin is dominated by the spacer refractive-index change,

$$\Delta n \approx \left[\frac{\partial n}{\partial T} \right] \Delta T \quad (2.17)$$

and from Wherrett, et al. [2.16] and Marburger, et al. [2.14]

$$\Delta T \cong \frac{\alpha I_{sp} r_0 l}{\kappa_s} \quad (2.18)$$

$$I_{sp} = I_1 \mathcal{T}(\lambda) \frac{(1-e^{-\alpha l})(1+R-e^{-\alpha l})}{\alpha l(1-R)e^{-\alpha l}} \quad (2.19a)$$

$$= I_1 \mathcal{T}(\lambda) \frac{1+R}{1-R} \quad \text{if } \alpha l \ll 1 \quad (2.19b)$$

and

$$\mathcal{T}(\lambda) = \frac{(1-R)^2 e^{-\alpha l}}{(1-Re^{-\alpha l})^2} \frac{1}{1+F \sin^2[\beta(I_{sp})]} \quad (2.20)$$

where ΔT is the temperature increase; I_{sp} is the mean cavity intensity and is greatest in the spacer layer; $\mathcal{T}(\lambda)$ is the transmission of a ZnS IF with linear absorption in the spacer layers being considered; $F \equiv 4Re^{-\alpha l}/(1-Re^{-\alpha l})^2$; α is the linear absorption coefficient of the spacer material; r_0 is the radius of a spot on the filter; and κ_s is the substrate thermal conductivity. The maximum temperature increase $(\Delta T)_{\max}$ occurs when the ZnS IF is on resonance, i.e., Eq. (2.20) becomes Eq. (2-11). Thus,

$$\Delta(T)_{\max} = \frac{P_I}{\kappa_s \pi r_0} \frac{(1-e^{-\alpha l})(1+Re^{-\alpha l})(1-R)}{(1-Re^{-\alpha l})^2} \quad (2.21a)$$

$$= \frac{P_I \alpha l}{\kappa_s \pi r_0} \frac{1+R}{1-R} \quad \text{if } \alpha l \ll 1 \quad (2.21b)$$

Assuming an input power of 20 mW, $R = 0.9444$, $\kappa_s = 0.01 \text{ W/cm} \cdot \text{K}^\circ$, and $l = 0.1-0.7 \mu\text{m}$, we have $\Delta(T)_{\max} = 10-80 \text{ C}^\circ$ for $r_0 = 25 \mu\text{m}$ and $\Delta(T)_{\max} = 60-400 \text{ C}^\circ$ for $r_0 = 5 \mu\text{m}$. Note the temperature increase $\Delta(T)_{\max}$ scales linearly with absorption coefficient α ; therefore, thermal load will be larger in the ZnSe case if other conditions are the same.

At normal incidence the critical input intensity I_c , the critical initial detuning $\beta_{0,c}$ and the critical input power P_c can be expressed as follows [2.5, 2.16].

$$I_c = \frac{\lambda \alpha}{2\pi n_2} f(R, \alpha l) \cong \frac{\lambda \alpha \kappa_s}{2\pi(\partial n/\partial T) r_0} \left[\frac{f(R, \alpha l)}{\alpha l} \right] \quad (2.22)$$

$$\beta_{0,c} = \frac{\sqrt{2}}{2} \frac{3(F+2)-s}{[(F+2)s-(F+2)^2-2F^2]^{1/2}} + 2\sin^{-1} \left[\frac{\sqrt{3F+2-s}}{2\sqrt{F}} \right] \quad (2.23a)$$

$$\cong \left[\frac{12}{F} \right]^{1/2} = \sqrt{3} \frac{1-Re^{-\alpha l}}{\sqrt{Re^{-\alpha l}}} \quad \text{in the high-finesse limit,} \quad (2.23b)$$

and

$$P_c = \pi r_0^2 I_c \cong \frac{\lambda \alpha \kappa_s r_0}{2(\partial n / \partial T)} \left[\frac{f(R, \alpha l)}{\alpha l} \right] \quad (2.24)$$

where f is a cavity factor defined as

$$f(R, \alpha l) = \frac{\sqrt{2}}{16} \frac{(1-R-\alpha l)^2}{(1-R)(1+Re^{-\alpha l})(1-e^{-\alpha l})} \frac{(3(F+2)-s)^2}{[(F+2)s-(F+2)^2-2F^2]^{1/2}} \quad (2.25)$$

and $s \equiv [(3F+2)^2-8F]^{1/2}$.

For a tilt angle θ in the spacer, l is to be replaced by $l/\cos\theta$ in all expressions and the mirror reflectance R calculated for that angle. From Eq. (2.25), the minimum critical input power scales linearly with wavelength, *thermal conductivity of a substrate*, *spot size*, and absorption coefficient, but scales inversely with $(\partial n / \partial T)$. The use of tight focusing (with a corresponding small spot size) and a low-thermal-conductivity substrate for the reduction of the power level have been proved effective experimentally. It is switching power rather than switching intensity that puts constraints on optical computing because there is limited power available from a laser and because one must remove the heat dissipated in the active device area. The critical input power of a ZnS IF, therefore, should be designed to be the minimum. This requirement implies that the following two conditions should be satisfied:

(i) The normalized cavity factor $f(R, \alpha l)/\alpha l$ should be minimized. For given coatings, there is an optimum value of αl for a minimum $f/\alpha l$. Referring to Eq. (2.25), it can be shown that in the high-finesse limit the minimum $f/\alpha l$ occurs when $\alpha l \cong 2(1-R)$ and $f/\alpha l \cong 2.6$ as depicted in Fig. 2.4. The minimum $f/\alpha l$ values are independent of R , i.e., the addition of more layers in high-reflectance mirrors will not further reduce the critical input power. It is evident from Fig. 2.4 that higher-reflectance stacks should be used to help reduce the experimental switching power. However, because α is small in the ZnS IF case, for the minimum $f/\alpha l$ to occur a long cavity length l is usually required, even when high-reflectance mirrors are used. For example, to obtain a minimum $f/\alpha l$ for the mirror stacks of order $p = 5$, a spacer thickness order $m \cong 26$ is required; this requirement is impractical in terms of optical transmission and real fabrication. Therefore, an attempt to reduce the spacer thickness by increasing the reflectance R is only marginally effective in a ZnS IF. In contrast, in a ZnSe IF, where α is larger, a spacer with thickness order $m \cong 4$ and mirror stacks of order $p = 5$ can be designed for ZnSe IFs (in which H:ZnSe $n \cong 2.722$ and $\alpha \cong 600 \text{ cm}^{-1}$; L:ThF₄ $n \cong 1.515$ and $\alpha \cong 0 \text{ cm}^{-1}$, all referred at $\lambda = 514.5 \text{ nm}$, are assumed) to obtain the minimum $f/\alpha l$. However, this design is

good only for reflection devices, because the ZnSe IFs with the optimization design in switching power will result in low transmission. This finding is supported by almost all ZnSe IFs reported so far in optical computing. Figure 2.5 shows the cavity factor f versus the number of spacer layers for various mirror stacks of order p , revealing that a mirror stack with a higher p will have a lower corresponding cavity factor f , and thus a tendency for a lower switching power. This result can be explained as follows. Mirrors having more (HL) stacks should have larger fields in the cavity; there is more absorption in the cavity, therefore, and less switching power is needed.

(ii) The material factor $\alpha/(\partial n/\partial T)$ should be maximized. Near the band edge the nonlinear index is the sum of the temperature dependence of the edge itself and a background thermo-optic coefficient:

$$\frac{\partial n}{\partial T} = \frac{\partial n}{\partial E_g} \frac{\partial E_g}{\partial T} + \left(\frac{\partial n}{\partial T} \right)_b \quad (2.26)$$

The analysis by Wherrett, et al. [2.16] indicates that in ZnSe material, under optimum cavity conditions, a large reduction in P_c near the band edge may not be obtained in practice, because $\alpha/(\partial n/\partial T)$ does not decrease at the edge. However, the optimization of the operating wavelength strongly depends on the precise form of the band edge, absorption values, and their temperature dependences. These data are not available in the literature, so a detailed experimental study is called for. The same is true for the ZnS case. The absorption change of a ZnS IF under different intensities (or under different temperatures) should be measured, e.g., by the method of probe-and-pump experiments [2.17] with an appropriate pump source, followed by application of a Kramers-Kronig transformation to obtain $\frac{\partial n}{\partial T}$, and deduction of the temperature increase ΔT in the spacer.

In brief, the ZnS IF design should first optimize the operating wavelength which depends on the availability of intense laser lines; this determines the absorption coefficient. Choice of the number of layers for high-reflectance mirrors then determines the spacer thickness that should be used; therefore a close thickness match in units of quarter-wavelengths between the mirror layers and the spacer layers should be chosen for the best results. However, we should note there is a tradeoff between low power consumption and high fan-out, because low power consumption requires a significant amount of absorption, while high fan-out requires a large transmission (i.e., low absorption). Similarly, a high contrast can be obtained only at the expense of high power consumption, as can be seen in Fig. 2.2. A good contrast

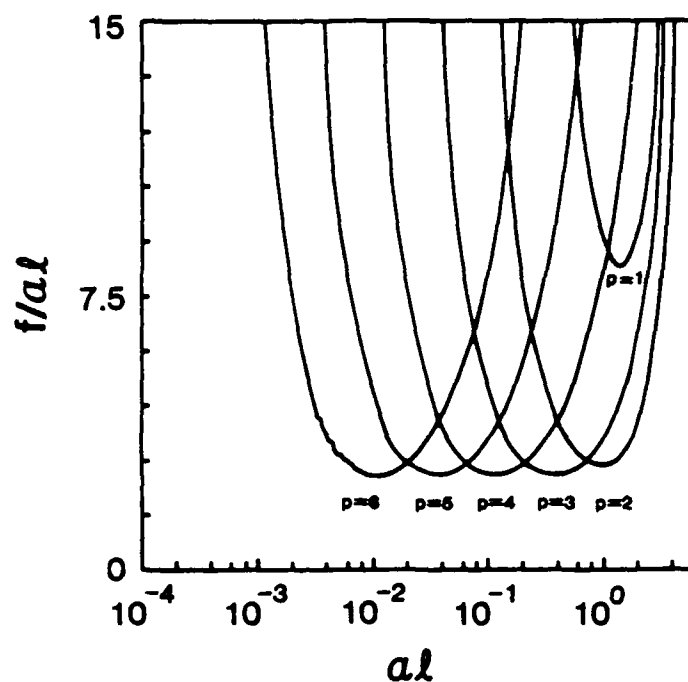


Figure 2.4 Normalized cavity factor $f/\alpha l$ for ZnS IFs; p is the mirror stack order.

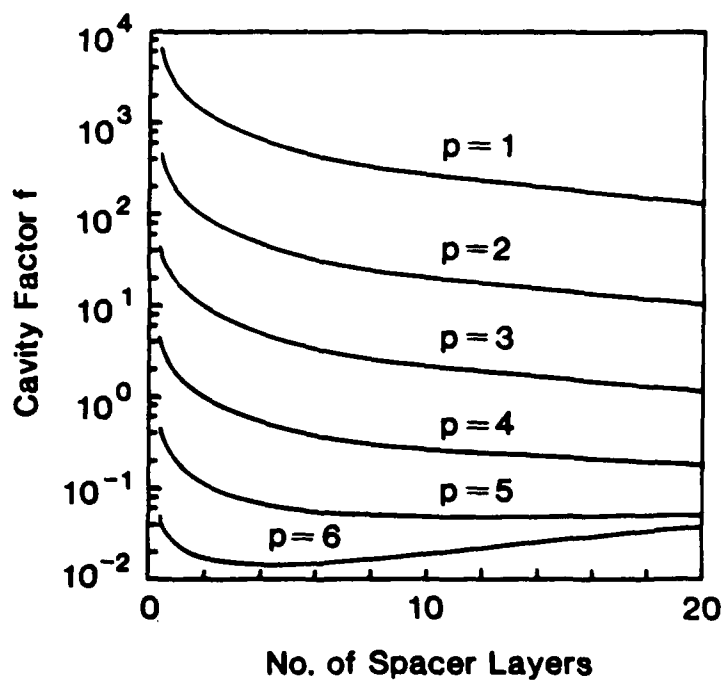


Figure 2.5 Cavity factor f vs numbers of spacer layers; p is the mirror stack order.

is obtained when the line A_1B_1 is close to the line A_2B_2 , where the bistable loop width becomes narrow.

We then consider the switch-up and switch-down times for a ZnS IF. Assume the cavity of an isolated spot in a ZnS IF initially is off resonance and is switched on resonance by switching on a mean incident power P_I . Heat conduction is assumed to be negligible. Switch-up time (τ_I) can be expressed, from thermal equilibrium, as follows.

$$\tau_I = \frac{\rho C_p V (\Delta T)_{\max}}{P_I A_I} \quad (2.27a)$$

$$= \frac{\rho C_p r_0 l^2 \alpha}{A_I \kappa_s} \frac{1+R}{1-R} \quad (2.27b)$$

where ρ is the density (5.3 g/cm³ for ZnS); C_p is the specific heat (0.339 J/g·C°); V is the volume of material being irradiated ($V = \pi r_0^2 l$ for a cylindrical volume); P_I is the incident power; and A_I the effective fraction of incident power being absorbed. Assume $A_I = 0.25$; then we have $P_I \tau_I \cong 10^{-6}$ – 10^{-8} J for a beam spot size of 20 μ m and $P_I \tau_I \cong 10^{-7}$ – 10^{-9} J for 10 μ m. These values are the physical limits on the switching energy in ZnS IFs. Correspondingly, if $P_I = 20$ mW, then the physical limits of switching time are $\tau_I \cong 400$ ns–16 μ s for a beam spot size ($= 2r_0$) of 20 μ m and $\tau_I \cong 200$ ns–8 μ s for 10 μ m. These values indicate that a thermal process is *not* necessarily slow if enough heat sink and adequate isolation of each spot are provided.

Heat conduction is the physical mechanism that determines the switch-down time (τ_d) of a thermally bistable ZnS IF. If a laser beam suddenly heats a cylinder of length much greater than the beam radius r_0 , solution of the heat diffusion equation leads to a thermal conduction of

$$\tau_d = \frac{r_0^2 \rho C_p}{(2.40)^2 K} \quad (2.28)$$

where K is the thermal conduction (0.167 W/cm C°). This relationship assumes that thermal properties are determined by transverse conduction. Equation (2.28) yields switch-down times of $\tau_d \cong 2$ μ s and 500 ns for spot sizes of 20 μ m and 10 μ m, respectively.

In our experiment, typical observed data are $\tau_I \cong 10$ –1000 μ s and $\tau_d \cong 20$ –2000 μ s.

It is apparent that fast switching-up and switching-down can be obtained by reducing the irradiated volume. However, for a given cavity factor there is a tradeoff between the switch-up time and the input switching power. Because density ρ and absorption A_f are directly related to the structure of a thin film, as the porosity increases the density decreases and the absorption increases. The porosity also reduces the heat flow through the film. This gives one possible explanation why a degraded ZnS IF showed a faster switching time but with lower transmission compared with a normal one.

2.4 Characterization and Results

Several ZnS IFs with $G|(HL)^p m(HH) (LH)^p |g|G$ structures were fabricated, where $m = 1, \dots, 6$ and $p = 4$ or 5 . Discussions will focus on the ZnS IFs with $p = 4$ only, because they constitute the majority. Although limited data are available, IFs with $p = 5$ do show lower linear peak transmission and lower switch-up power as expected from Table 2.1 and Fig. 2.5, respectively.

The linear characteristics with respect to the peak transmission and bandwidth (FWHM) of ZnS IFs are determined with a Cary-14 spectrometer. Uniformity of these linear characteristics is determined by sampling five different locations on a ZnS IF. These locations are defined by a paper mask as depicted on the upper-right corner in Fig. 2.6, which is attached to the substrate side of a ZnS IF. A typical measurement result is also shown in Fig. 2.6. The variation of peak wavelength, peak transmission and FWHM can provide information about the magnitude of the layer thickness error, the difference of reflectance between two mirrors, the amount of loss (scattering), etc. Some of our ZnS IFs have zero peak-wavelength shift (reference wavelength = 514 nm) over a large area as depicted in Fig. 2.7. In Fig. 2.8, linear peak transmission is plotted as a function of the spacer thickness for the calculated values and for the experimental data. Figure 2.9 shows that the calculated critical initial detunings, the calculated FWHMs, and the experimental FWHM data all vary as a function of the spacer thickness. The critical initial detuning is calculated on the basis of Eq. (2.23a) with the transformation relationship $\Delta\lambda_{\beta_{0,c}} = \lambda_0 \beta_{0,c} / 2\pi(m+p)$, and in a high-finesse limit the critical initial detuning is related to the FWHMs in Eq. (2.12) by $\Delta\lambda_{\beta_{0,c}} \cong \sqrt{3}/2$ FWHMs. The calculated FWHMs are from Table 2.1 and could be regarded as the minimum initial detuning required for optical bistability as derived in Eq. (2.9).

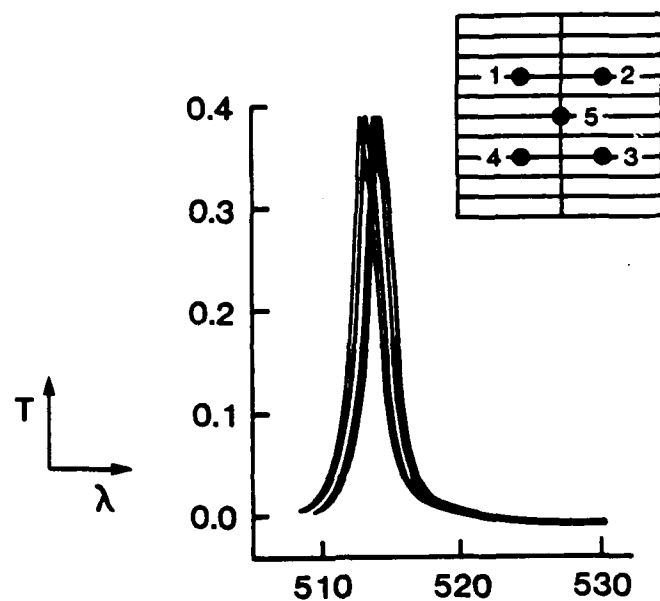


Figure 2.6 An example of transmission measurements on a ZnS IF. The upper right corner is a mask used to measure the uniformity; the center-to-center distance between spots 1 and 2 is 21 mm.

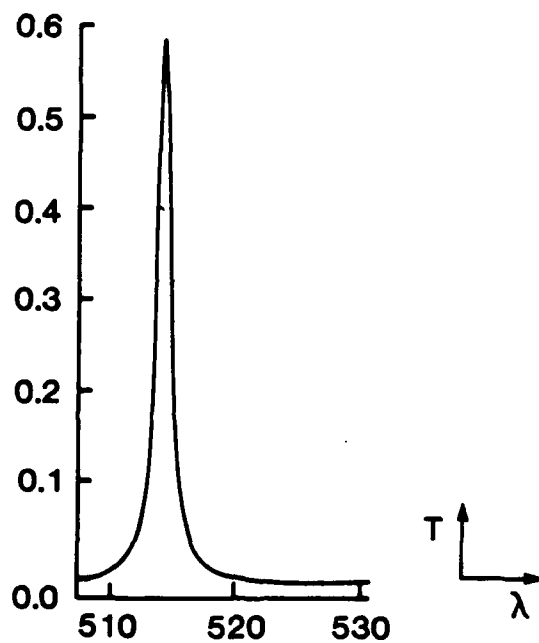


Figure 2.7 An example of a fairly uniform ZnS IF.

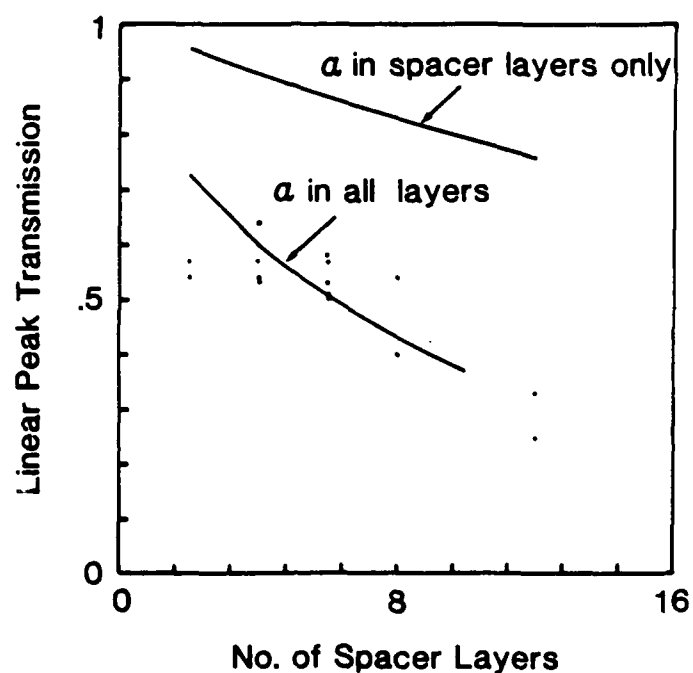


Figure 2.8 Linear peak-transmission of a ZnS IF vs numbers of spacer layers. Solid lines are of theoretical calculations; dots are experimental data of IFs with $p=4$.

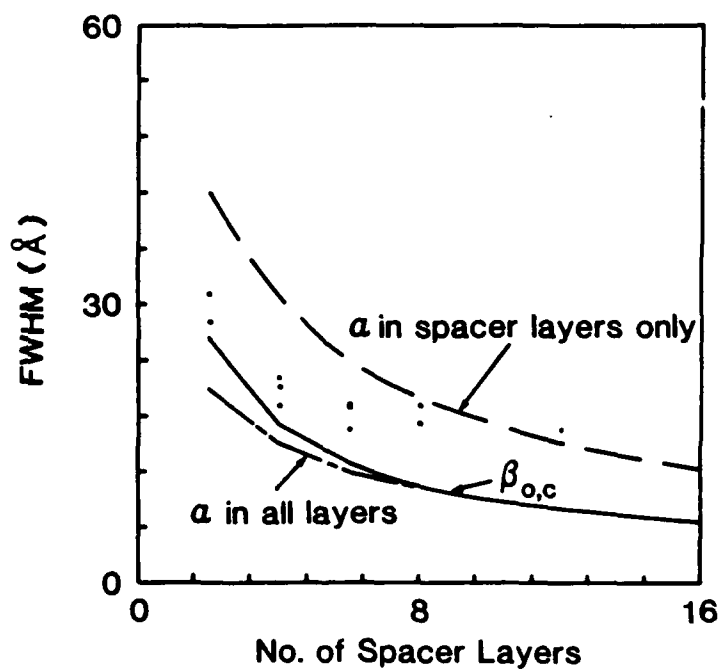


Figure 2.9 FWHMs of a ZnS IF vs numbers of spacer layers. Solid lines are calculated critical initial detuning $\beta_{0,c}$ and calculated FWHMs (see text); dots are experimental data of IFs with $p=4$.

To characterize the nonlinear properties of ZnS IFs, the cw output from a Coherent Innova 90-4 Ar⁺ laser operating at 514 nm is used. The beam is modulated by an assembly of a rotating half-wave plate and a polarized beam splitter with a total noise level of about 3%. An IF is put at the center of a rotational stage so that tilt angles can be read, and the whole setup is mounted on a X-Y-Z translation stage to have fine adjustment. A lens is used to provide a beam spot size on an IF of roughly 20 μm measured by a pinhole set. Two pin-photodiodes (United Detector Technology) are used to detect input and output signals of an IF, and these signals are sent to a Hitachi V-1070 100-MHz oscilloscope to trace a bistable loop. In all power measurements, the substrate side of a ZnS IF always faces the incident beam. A Coherent Model 212 power meter is used to determine the minimum switching power with a calibration factor of 2 for $\lambda = 514.5 \text{ nm}$.

Figure 2.10 shows the experimental data of the minimum switch-up power and the calculated values versus spacer thickness. The calculated values are based on Eq. (2.24), with $k_s = 0.01 \text{ W/cm}\cdot\text{K}^0$, $r_0 = 10 \mu\text{m}$, and $\partial n/\partial T \cong 8.02 \times 10^{-5} \text{ K}^{-1}$ [2.18]. The solid line is for equal R; the dashed line is for unequal R with 10% less R for the front mirror and illumination from the front; the dash-dotted line has the same unequal R as the dashed line, but illumination is from the substrate side. It is interesting that the reflectance mismatch of two mirrors can be clearly identified by comparing the difference of the minimum switch-up powers obtained when the incident beam is from the front and from the rear. The calculated data show that the thicker the spacer the lower the switch-up power, and that for a fixed reflectance the minimum switch-up power decreases abruptly for *the first few layers*, indicating that the spacer thickness is an important factor in the switching power optimization. Experimental data of the switch-up power, though scattered, do show this trend except that these data are consistently smaller than the calculated ones. The deviation may result from errors in the parameter values used in Eq. (2.24) for the calculation, and from errors attributable to the simplified assumptions that have been made (e.g., only the spacer layers include linear absorption, and the temperature through the multilayer is assumed to be uniform). Equation (2.24) predicts that the use of a low thermal conduction substrate will reduce the value of the minimum switching power. This is proved experimentally by a reduction of 2-3 times in switch-up power for a ZnS IF with the same $(\text{HL})^4 4\text{H}(\text{LH})^4$ structure, but with a 5-mm-thick SF glass substrate ($\kappa_s \cong 0.001 \text{ W/cm}\cdot\text{K}^0$), instead of a 1-mm-thick Kodak glass substrate ($\kappa_s \cong 0.01 \text{ W/cm}\cdot\text{K}^0$). Figure 2.11 shows that the linear dependence of the switch-up power on the beam spot-size calculated from Eq. (2.24) is consistent

with the linear trend of observed data. Objectives with powers ranging from N.A. = 0.1 to N.A. = 0.3, which are assumed to have diffraction-limited performance, are used to provide different spot sizes. A ZnS IF with the structure (HL)⁴4H(LH)⁴ is chosen arbitrarily as an example. For higher N.A. an irreversible deterioration of bistable characteristics has always been observed with continuing illumination. It is suspected that the deterioration comes from the relatively high temperature in the spacer which perhaps exceeds the damage threshold of the spacer material. Similar results are reported by Sahlen, et al. [2.19]. This imposes an inherent difficulty in implementing multiple switches on a ZnS IF for optical computing. Simultaneous switching of 8 pixels with separation of $\approx 100 \mu\text{m}$ on a ZnS IF has been observed by incorporating a 1-D binary phase grating with a lens (focal length = 50 mm). The IFs are used later in the three-spot pattern recognition experiment described in Section 13.

The fan-out of a bistable device can be defined as the difference in transmitted power before and after the switch-up divided by the signal input power required to switch-up when the device is biased. The contrast can be defined as the ratio of the upper-branch transmission of a holding beam just after switch-on to the lower-branch transmission just before switch-on. In the gating experiments the holding beam is always modulated synchronously with the signal beam instead of cw operation. This scheme has two advantages: 1) it reduces the heat load in the devices without reducing the data rate; 2) in the optical bistability mode the bias can be adjusted as close to the switch-on point as noise will allow, therefore a better fan-out is achieved.

We have obtained a fan-out of 4 using a ZnS interference filter. A 40-mW beam held the filter's operating point just below bistable switch-up; a 2-mW input then yielded an 8-mW change in an output power. In contrast, in electronics, TTL typically has fan-out of 10 and CMOS of 50. Cascading has also been performed by two ZnS IFs. A portion of the output beam from the first IF and a holding beam are used to switch the next IF. The largest contrast we have observed in ZnS IFs is about 4, while for ZnSe it is about 3. This may be due to the higher transmission of ZnS filters, which gives them an additional advantage of a larger absolute output. As discussed before, ZnSe IFs differ from ZnS IFs in the respect that ZnSe IFs have a higher absorption, and $(1/n) \cdot (\partial n / \partial T)$ in bulk ZnSe is about twice that for bulk ZnS near the absorption band edge. Consequently, it may be advantageous to use ZnSe IFs for low-power applications (e.g., in the last stage of a system). ZnS IFs could be used to switch a number of pixels on a ZnSe IF when a large fan-out is needed or when the noise level requires a high absolute switching signal.

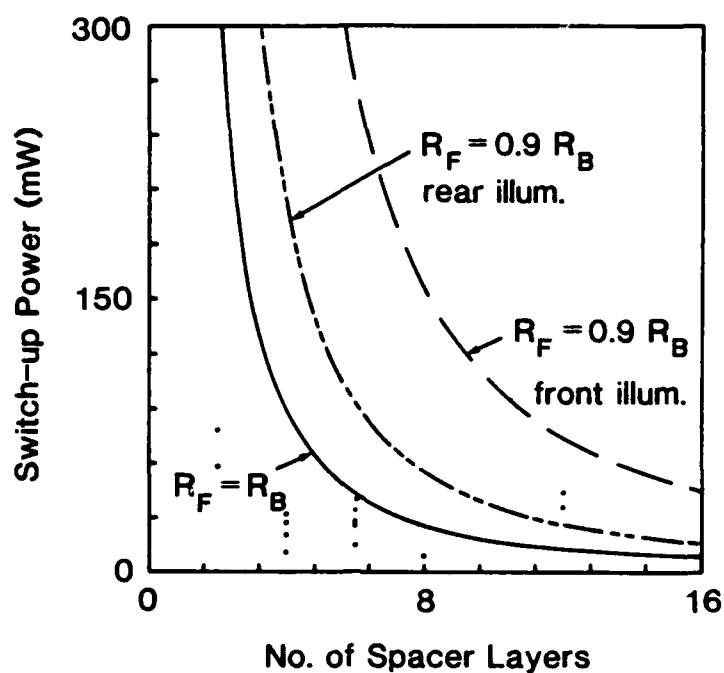


Figure 2.10. Minimum switch-up power of a ZnS IF vs numbers of spacer layers. Solid, dashed, dash-dotted curves are calculated value (see text); dots are experimental data of IFs with $p=4$.

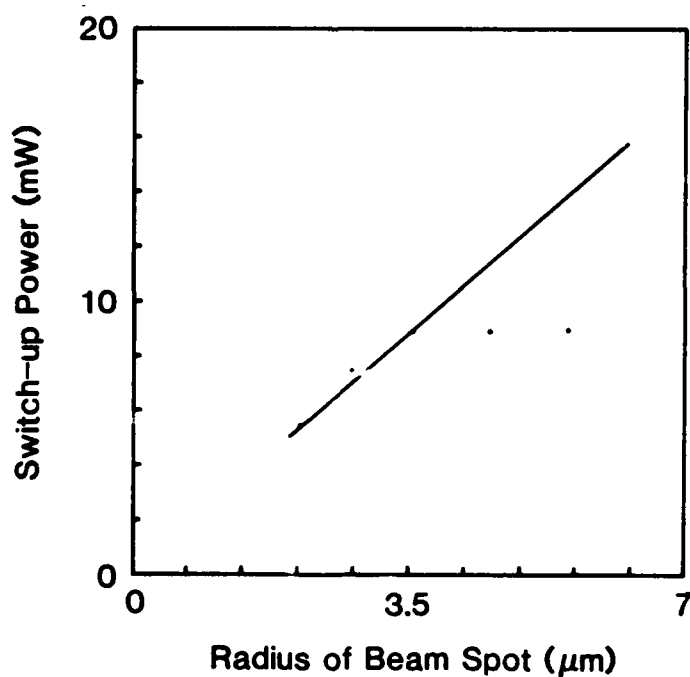


Figure 2.11 Switch-up power versus radius of different beam spot. Solid line for calculated data; dots for experimental data of IFs with $(HL)^4 4H(LH)^4$.

The use of transmission or reflection output from a bistable IF with an appropriate holding beam, as depicted in Fig. 2.12, makes the following logic functions available: AND, NAND, OR and NOR. For a logic AND operation, two units of light intensity are necessary for logic ONE; while for a logic OR operation, one unit of light intensity will be sufficient. In addition, non-latching logic gates can be realized with the proper control of detuning to reduce the bistable width to zero.

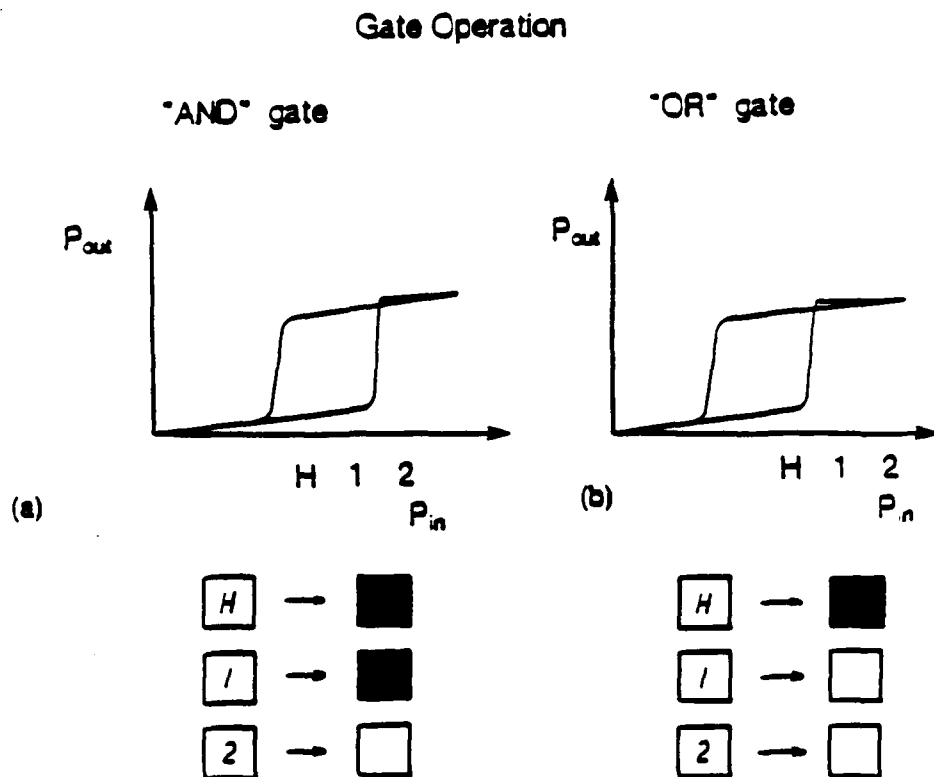


Figure 2.12 Different amounts of holding beam are used for (a) an AND gate operation and (b) an OR gate operation.

2.5 Discussion

In summary, ZnS IFs are relatively simple and inexpensive to grow with reasonable uniformity (less than 25% variation in the bistability switch-on power over $12 \times 21 \text{ mm}^2$ has been observed in our experiments). The nonlinearity arises from the shift of the band edge with heating, so it is only weakly resonant. Operation is good at 514.5 nm and at 488 nm, which permits the use of many-watt Ar⁺ lasers to produce multiple beams. In addition to facilitating alignment, visible light offers a

convenient means for learning to work with many beams in parallel, and is impressive during demonstrations of optical-computing concepts. ZnS filters do, however, possess undesirable features. Their thermal response (0.01 to 1 ms) is much slower than that of GaAs, but their power per pixel (10 mW) is about the same as for GaAs. A ZnS and/or ZnSe IF and an etalon of bulk GaAs and/or of GaAs/Al_{1-x}Ga_xAs multiple-quantum-well structure are two systems currently under intensive investigation for use as all-optical logic gates. Comparisons of these two systems are summarized in Table 2.2.

The packing density of logic gates is determined by the minimum interpixel separation required to prevent crosstalk. Crosstalk is an undesirable occurrence in which an element switched to the ON state can cause its neighbors to switch through some form of transverse coupling. In GaAs devices the dominant mechanism is carrier diffusion, while in IFs it is thermal diffusion. Although an interpixel separation of 20 μm was shown to prevent crosstalk in a two-beam experiment [2.9], results obtained by Abraham, et al. [2.20] suggest that the pixellization for optical logic array in a ZnS IF will be an effective way to achieve low crosstalk and low switching energy. Suppose that a 5 cm \times 5 cm bistable ZnS IF can have 10^6 operating spots, and that each spot requires 10 mW of power to be in ON state; one then could operate at a 10-kHz rate, resulting in 10^{10} bit operations per second. However, 10 kW of laser power would be required. It is a challenging thermal engineering problem to remove a heat load of 100 W/cm², assuming 25% absorption per spot.

Future work could be directed to pixellation and to innovative methods for reducing the power consumption per pixel. One example of reducing the switch-up power was demonstrated by using an extra-cavity absorber [2.21]. Although those devices operate only in the reflection mode, the operating wavelength has been extended to the near infrared, making the IFs compatible with laser diodes. Other fabrication techniques have been adapted to improve the IFs' properties. For example, observations by Chow, et al. [2.22] show that samples produced by MBE are likely to be inherently more stable because the spacer regions are dense and contain no porosity. Still, engineering problems in conventional IFs, such as how to obtain long-term stability against water-vapor shifts and how to remove heat load, remain to be solved.

Table 2.2 Comparison between Zn/ZnSe IFs and GaAs etalons for all-optical logic gates based on optical bistability.

	GaAs/Ga _x Al _{1-x} As etalons	ZnS/ZnSe IFs
refractive index	3.48-3.6 (bulk)	2.35-2.45
absorption coeff.	≈ 0.1 - $1.0 \mu\text{m}^{-1}$	100-600 cm^{-1}
bandgap	GaAs: 870 nm GaAs/GaAlAs: 806-870 nm	≈ 300 nm
mechanism	electronic refractive	thermal refractive
$\lambda_{\text{operation}}$	GaAs: 885 nm GaAs/GaAlAs: tunable 820-885 nm	514 nm, 488 nm
structure	bulk, MQW	G(HL) ^P mHH(LH) ^P G
substrate	GaAs	glass
$T_{\text{peak,linear}}$	10-20%	$\approx 60\%$ for ZnS
FWHM (Å)	≈ 20	≈ 15
spot size (μm)	≈ 15	≈ 20
$P_{\text{min.}}$ (mW)	≈ 10	≈ 10
switch-up time	$\approx \text{few ps}$ windowless GaAs: 30 ps	10-1000 μs
2-D array realized?	Yes	No
contrast	10:1	3-4:1
fan-out	4	4
uniformity	Good	Good
single λ operation	Yes	Yes
gate operation	OR NOR AND NAND	OR NOR AND NAND
cascading?	Yes	Yes
compatible with IC	No	No
compatible with laser diode?	Yes	No

3. OPTICAL NONLINEARITY OF GaAs AND GaAs-AlGaAs MULTIPLE-QUANTUM WELLS

3.1 Introduction

Semiconductors have been studied extensively for possible applications to fast opto-electronic devices, since they exhibit large optical nonlinearities in the vicinity of their band edge [3.1, 3.2]. The microscopic origins of the electronic nonlinearities in semiconductors are explained by many-body effects such as screening of Coulomb interactions, bandgap renormalization and phase-space filling [3.3, 3.4]. Recently, it has been reported that room-temperature optical nonlinearities in bulk GaAs are attributable to plasma screening of the Coulomb enhancement and to band filling [3.5].

Modern crystal growth techniques like MBE and metalorganic chemical vapor deposition (MOCVD) make possible high quality MQW structures with precise layer thicknesses and high uniformity. The confinement of electron-hole pairs in MQWs enhance their excitonic effects and makes them observable even at room temperature [3.4, 3.6]. The magnitude of the nonlinear refractive index has been measured for two GaAs/AlGaAs MQWs using degenerate four wave mixing [3.4]. Generally, it has been believed that MQWs have a larger nonlinear refractive index than does bulk GaAs. However, there has been no systematic study of the effect of well size on the optical nonlinearities in MQWs.

We have made measurements of the nonlinear refractive indices of GaAs/AlGaAs MQWs as a function of well size, and compared them with those of bulk GaAs. The nonlinear absorption spectra for 76-Å, 152-Å, and 299-Å GaAs/AlGaAs MQWs and for bulk GaAs were measured. The changes in the refractive index were obtained from a Kramers-Kronig transformation of the measured absorption changes. The index variation Δn , and the ratio of maximum Δn to the carrier concentration N , $\Delta n/N$, were compared for these samples.

3.2 Experimental Results

A nitrogen-pumped dye laser operating at 816 nm was used for the pump beam. The probe consisted of the spectral broadband spontaneous luminescence emitted by a cell containing IR-144 or HITC dyes. The nitrogen laser synchronously pumped both the dye laser and the dye cell which subsequently delivered pulses of 3 ns full width at half maximum (FWHM). The two beams were collimated and

brought onto the sample so as to be coincident in time and space. This was carefully checked using an oscilloscope and TV a monitor. The spot diameter of the pump beam at the sample was $\approx 200 \mu\text{m}$, allowing us to neglect the diffusive decay time in comparison to the recombination time [3.4]. The probe beam diameter was $\approx 50 \mu\text{m}$. The pump beam was polarized orthogonal to the probe beam, which permitted us to block the pump beam selectively before it entered the spectrometer. The transmission of the sample for the broadband probe beam was measured using a spectrometer equipped with an optical multichannel analyzer (OMA).

Four different MBE-grown samples were used for this research: 76-Å, 152-Å and 299-Å GaAs/AlGaAs MQWs and bulk GaAs. The first MQW sample consisted of 63 layers of 76-Å-thick GaAs. Each GaAs layer is followed by an 81-Å-thick layer of $\text{Al}_{0.37}\text{Ga}_{0.63}\text{As}$. The total GaAs thickness of this sample was $0.48 \mu\text{m}$. The next sample consisted of 100 layers of 152-Å-thick GaAs and 104-Å-thick $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ barriers for a total GaAs thickness of $1.52 \mu\text{m}$. The third sample had 61 layers of 299-Å-thick GaAs and 98-Å-thick $\text{Al}_{0.36}\text{Ga}_{0.64}\text{As}$, resulting in a total GaAs thickness of $1.8 \mu\text{m}$. The bulk GaAs consisted of a $2.05\text{-}\mu\text{m}$ GaAs layer sandwiched between two AlGaAs layers. These four samples are called 76-Å, 152-Å, 299-Å and bulk, respectively, throughout this section. Different total GaAs thicknesses for the four samples does not affect either the measured absorption changes or the nonlinear index, since the thickness factor drops out when the absorption coefficient is obtained from the probe transmission.

The absorption changes for the various pump intensities were measured in the vicinity of the band edge at room temperature. The pump wavelength was fixed above the band edge at 816 nm. The same pump beam intensities were used for all four samples. Figure 3.1(a) shows the nonlinear absorption spectra for bulk GaAs. A very small exciton feature appears in the linear spectrum labeled 0. Exciton screening is observed for small pump intensities, while broad absorption changes in the band become more pronounced as the pump intensity is increased. The changes in the refractive index were obtained from a Kramers-Kronig transformation of the changes in the measured absorption spectra. Figure 3.2(a) shows the resulting variation in the refractive index. The rising edge observed for high energies is an artifact from the use of a finite integration interval in the Kramers-Kronig transformation [3.7]. The error introduced by clipping this interval is less than five percent in the region of interest near the exciton resonance. This was verified by numerically integrating the absorption changes extended over an infinite frequency

range and comparing this result with the Δn computed using finite integration limits [3.7]. The validity of this indirect technique of measuring the nonlinear index has been confirmed by several direct interferometric measurements [3.5].

Excitons and carriers are confined to quasi- two dimensions in MQWs. This confinement causes a shift in the linear absorption edge to higher energies, enhanced excitonic features, and a splitting of the heavy- and light-hole excitons [3.4]. The exciton peak, barely resolvable in bulk, is more prominent in the 299-Å sample, as shown in Fig. 3.1(b). At low pump intensity, exciton saturation dominates. At high intensity, the nonlinear behavior is similar to that of bulk GaAs because of the relatively weak confinement in the 299-Å sample. In Fig. 3.1(c) and (d) are displayed similar curves for the 152-Å and 76-Å samples. The increase in the excitonic absorption now is very apparent. Note the clear separation of the heavy- and light-hole excitons in the 76-Å sample, in contrast to that of the 152-Å sample. The corresponding variation in refractive index is shown in Fig. 3.2(c) and (d).

3.3 Discussion of Results

Curves of Δn versus intensity for the four samples are plotted in Fig. 3.3(a). Note from Fig. 3.3(a) that to produce the same Δn , a lower pump intensity is required for the 76-Å MQW sample than for the bulk sample. These results indicate that the smaller MQWs have larger nonlinear refractive indices. However, one must note that the comparison of intensity-dependent index changes is somewhat misleading. For resonant excitation, changes of the optical material properties depend on the excitation intensity only through the carrier density. Moreover, the samples each have a different absorption at the pump frequency, so that a single excitation intensity generates a different number of carriers for each sample. Therefore, to obtain a clearer comparison of the material nonlinearities themselves, we compare the index change per excited carrier, $\Delta n/N$, for these samples.

Figure 3.3(b) shows the maximum $\Delta n/N$ for a given carrier concentration. The carrier concentration is obtained using the simple rate equation:

$$\frac{dN}{dt} = \frac{\alpha(\omega_{\text{pump}})I}{\hbar\omega_{\text{pump}}} - \frac{N}{\tau} \quad (3.1)$$

where τ is the lifetime of the electron-hole pairs, ω_{pump} is the frequency of the pump beam and $\alpha(\omega_{\text{pump}})$ is the linear absorption coefficient measured at the pump frequency. Assuming a Gaussian temporal profile for the pump beam, Eq. (3.1) yields

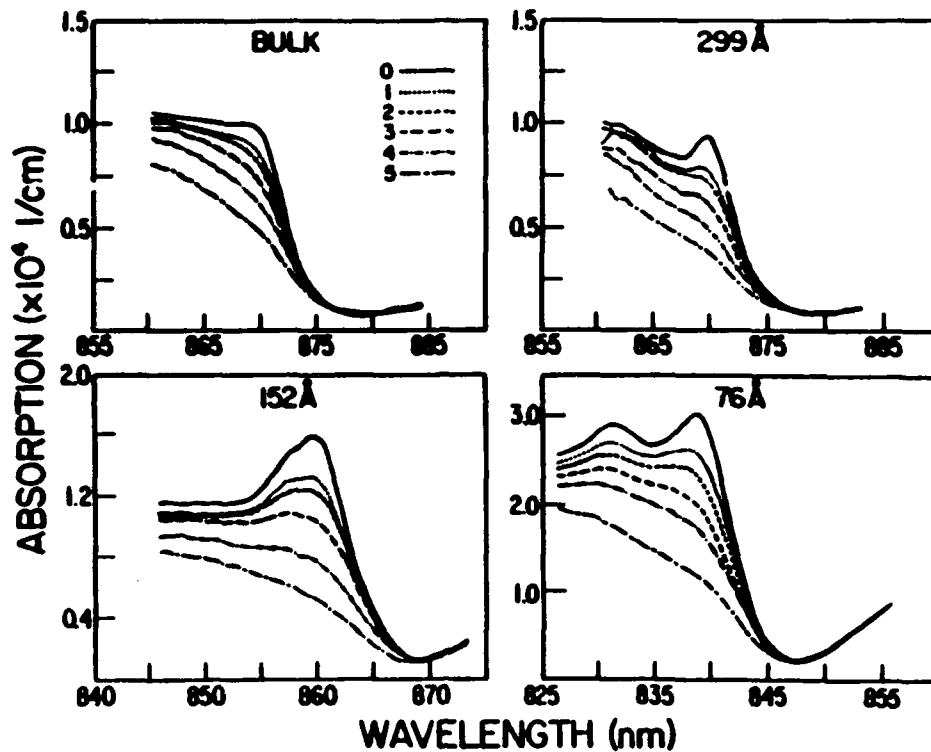


Figure 3.1. Experimental room-temperature absorption spectra for: (a) bulk GaAs; (b) 299-Å MQW; (c) 152-Å MQW; and (d) 76-Å MQW. The curves labeled in the figure represent pump beam intensities (in W/cm^2) of: (0) 0; (1) 670; (2) 1270; (3) 2650; (4) 5400; and (5) 11700.

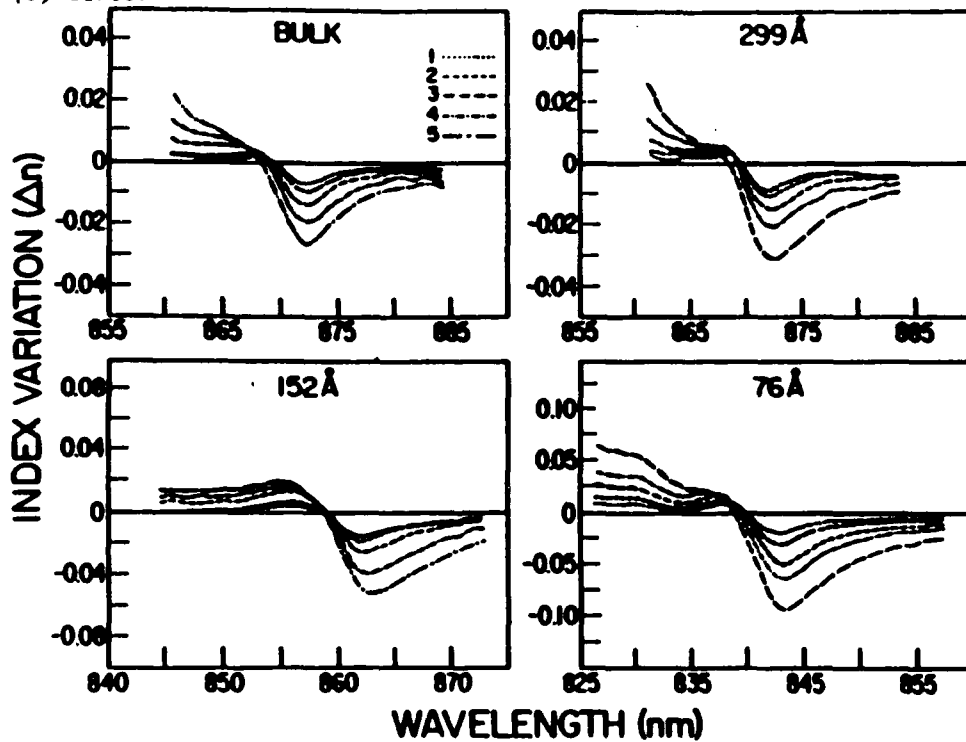


Figure 3.2. Refractive index changes corresponding directly to the measured changes in the absorption spectra for: (a) bulk GaAs; (b) 299-Å MQW; (c) 152-Å MQW; and (d) 76-Å MQW.

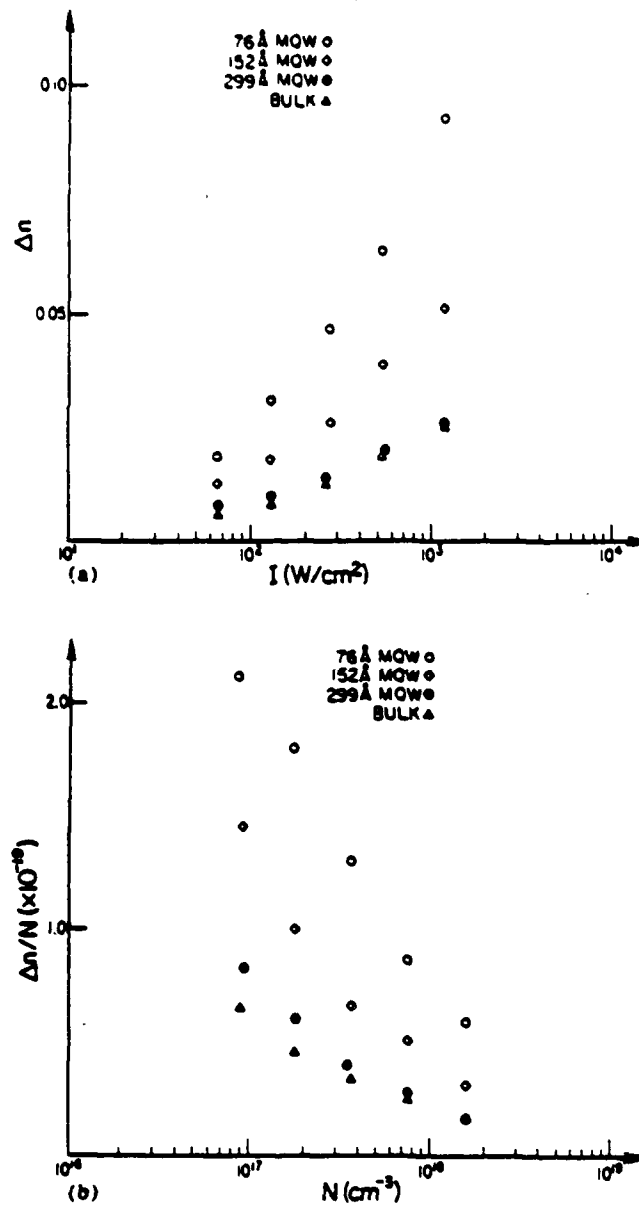


Figure 3.3 (a) Maximum change in refractive index vs. pump beam intensity; (b) maximum change in index per carrier concentration vs. the carrier concentration.

$$N(t) = \frac{e^{-t/\tau} \alpha(\omega_{\text{pump}}) I_0}{\hbar \omega_{\text{pump}}} \int_{-x}^t e^{t'/\tau} \exp \left\{ - \left[\frac{1.6651}{(\text{FWHM})} t' \right]^2 \right\} dt' \quad (3.2)$$

where I_0 is the peak pump intensity, x is chosen to be twice the FWHM, and t is the delay time between the pump and probe.

As shown in Fig. 3.3(b), $\Delta n/N$ increases by a maximum factor of 3 for $N \cong 10^{17} \text{ cm}^{-3}$ as the MQW well size decreases from bulk to 76 Å. The carrier

lifetime was also measured in these samples. Table 3.1 gives the measured values for the carrier lifetime.

Table 3.1 Carrier lifetime in bulk and MQW GaAs

Sample	Lifetime
1. Bulk (2.05 μm)	32 ns
2. 299-Å MQW	24 ns
3. 152-Å MQW	18 ns
4. 100-Å MQW	10 ns
5. 76-Å MQW	3 ns

As can be seen from Fig. 3.1, the excitons bleach and broaden with increasing carrier concentration. In contrast to bulk GaAs, [3.5] it is reported that exciton saturation due to phase space filling and exchange are the dominant nonlinear mechanisms for the MQWs [3.3, 3.8]. Several models were employed in order to determine the excitonic absorption saturation with carrier concentration for the MQWs. The best fit to the experimental data was obtained using the following saturation model:

$$\alpha(N) = \frac{\alpha_0}{1 + N/N_s} \quad (3.3)$$

where α_0 is the linear absorption coefficient at the heavy-hole exciton peak and N_s is the saturation carrier concentration. N_s and α_0 for the MQW samples, as determined from Eq. (3.3), are given in Table 3.2. The larger α_0 for the smaller-well MQWs results from the decrease in their exciton Bohr radii, a_B (because the transition probability is proportional to $1/a_B^3$). These results indicate that the saturation density N_s is nearly well-size independent and that the factor-of-3 increase in the optical nonlinearities are attributable mainly to the factor-of-3 increase in the exciton absorption.

Table 3.2 Values of α_0 and N_s determined by least squares fit.

Sample	$\alpha_0(\times 10^4 \text{ cm}^{-1})$	$N_s(\times 10^{17} \text{ cm}^{-3})$
76-Å MQW	2.7	9.0
152-Å MQW	1.4	7.2
299-Å MQW	0.8	11.1

The refractive nonlinearity increases as the well thickness is reduced to 76 Å. The optimum well size is probably around 50 Å, at least with present-day, state-of-the-art growth (narrower wells usually have broader lower exciton peaks, resulting from inhomogeneous broadening caused by fluctuations in thickness). As stated earlier, the number of layers and their thickness do not affect the measured nonlinearity for a sample. However, greater thicknesses correspond to greater nonlinear phase shifts for etalon devices ($\Delta\phi = \frac{2\pi}{\lambda} \Delta n l$ where λ is the wavelength, l is the thickness, Δn is the index change and $\Delta\phi$ is the phase shift). In addition, transmission decreases as layer thickness increases. Finally, the carrier lifetime is shorter for smaller well sizes, as shown in Table 3.1. These properties, greater index changes with shorter decay times, make smaller-well-size etalons more desirable for application in all-optical devices.

As the well size is reduced, the exciton peak is shifted to shorter wavelengths as a result of the confinement effect. For example, the exciton peak occurs at $\approx 870\text{nm}$ for a 299-Å MQW, while it occurs at $\approx 840\text{nm}$ for the 76-Å MQW.

The absorption of excitons in MQWs is polarization dependent, i.e., both heavy-hole and light-hole excitons are observed for the E field in the plane of the layers, while only the light hole remains for the E field perpendicular to the layers. Therefore, the absorption spectrum effectively is shifted to the blue range for light polarized perpendicular to the layers. We have already employed this polarization property of the MQWs in nonlinear directional couplers. However, most of our etalon experiments were performed with light polarized parallel to the plane of the MQWs, in which case polarization effects are negligible.

We have studied excitonic nonlinearities in MBE-grown ZnSe crystals, using techniques like those presented here for GaAs; see Appendix A.

3.4 Conclusion

We performed a systematic study of the dependence of the optical nonlinearities on well thickness for GaAs/AlGaAs MQWs (of 76-Å, 152-Å and 299-Å) and for bulk GaAs at room temperature. The measured ratio of the nonlinear refractive index to the carrier concentration increased by a factor of 3 at these carrier densities as the well size decreased from bulk to 76 Å. The increase in excitonic absorption for the smaller MQWs appears to be responsible for this behavior. We note that the larger nonlinearity and the better compatibility with existing semiconductor lasers tends to favor MQWs over bulk GaAs for device applications. Comparisons between bulk and MQW band edge nonlinearities are summarized in Reference [3.9].

4. UNIFORMITY OF GaAs ETALONS

When our Statement of Work was written, it was believed that the GaAs etalon worked well as a single device, and that most of our research should be devoted to multiple beams and to interconnecting etalons. The discovery of the gain problem for short pulses (see Section 6) prompted a reassessment, and we placed less emphasis on multiple-beam considerations. In addition, our principal source of GaAs samples, Art Gossard, left Bell Labs and went to the University of California at Santa Barbara, where he found that the MBE machines there had been acquired by people interested in competing rather than in collaborating with us. This further hampered our uniformity studies.

We performed two systematic studies of uniformity as reported here, one with external mirrors, and one with MBE-grown mirrors.

4.1 External Mirrors [4.1]

The nonuniformity of standard GaAs etalons results from: 1) nonuniformity of the MBE growth; 2) nonuniformity introduced during etching of the substrate; and 3) increased or decreased nonuniformity resulting from the attachment of external mirrors to the GaAs crystal.

That MBE growth often has appreciable nonuniformity is shown in Fig. 4.1, where the peak-to-peak variation over 0.06 mm^2 is $0.2 \lambda/n_0$, with $\lambda = 850 \text{ nm}$ and $n_0 = 3.6$, and an RMS variation of $0.03 \lambda/n_0$. In general, etching away the substrate results in greater nonuniformity, a wedge is introduced during the etching process, and persists in the stop layer (but is diminished by the approximate 30:1 difference in etch rates). This wedge nonuniformity can be compensated for by adjusting the external mirrors. A brute-force method is to put the sample in an external Fabry-Perot etalon, as shown in Fig. 4.2.

The layout for testing the samples is depicted in Fig. 4.3. The dye laser is calibrated by connecting a fiber to a monochromator with an optical multichannel analyzer. The pulse width is $0.8 \mu\text{s}$ and the duty cycle is $0.07 \mu\text{s}$. Neutral density (ND) filters are used to make sure there is no nonlinear effect. By tuning the dye laser, the Fabry-Perot transmission profiles are obtained. On the monitor, the fringes on the sample can be seen easily. These fringes are the result of nonuniformity of the sample, evidenced by the fact that the fringes move when the sample is moved. Through an adjustment of the screws on the holder, the number of fringes on the sample can be reduced to very few, indicating that the wedge has been compensated for.

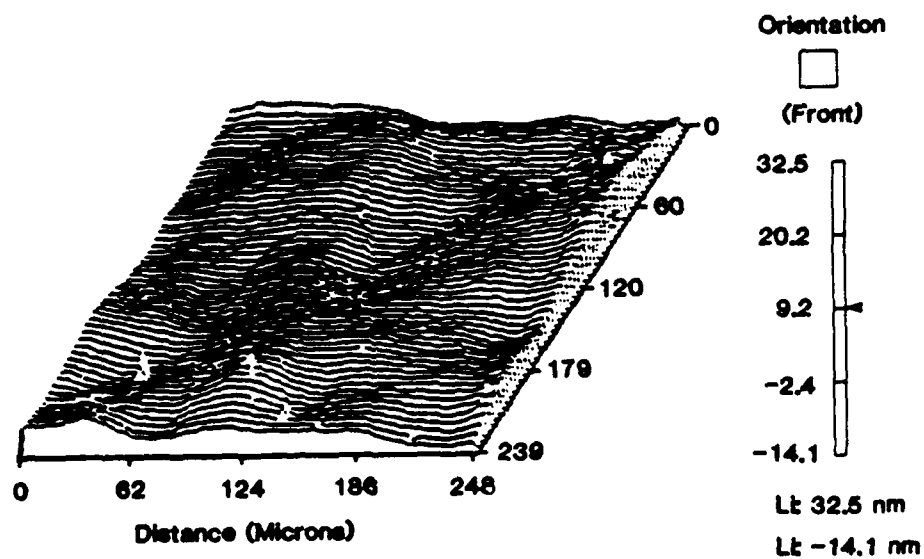
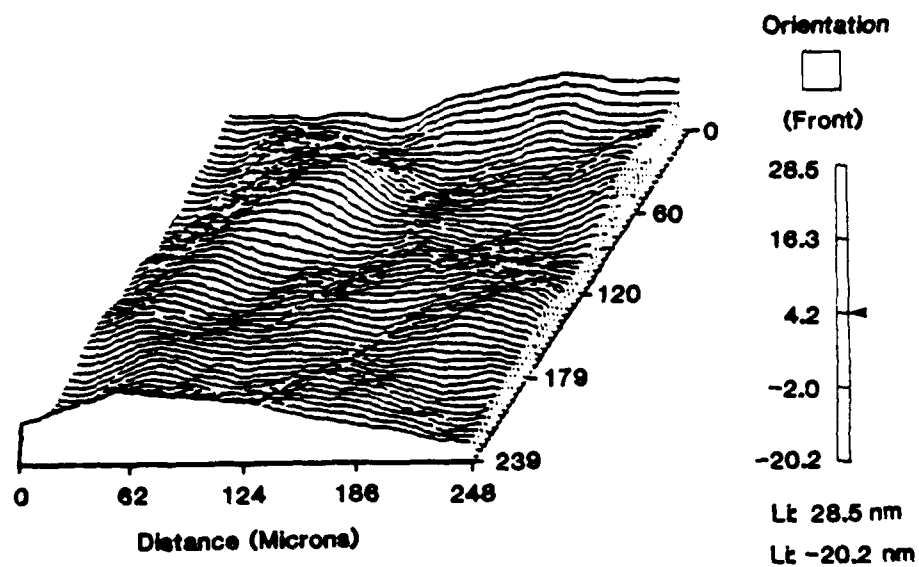


Figure 4.1 Surface nonuniformity of MBE-grown sample.

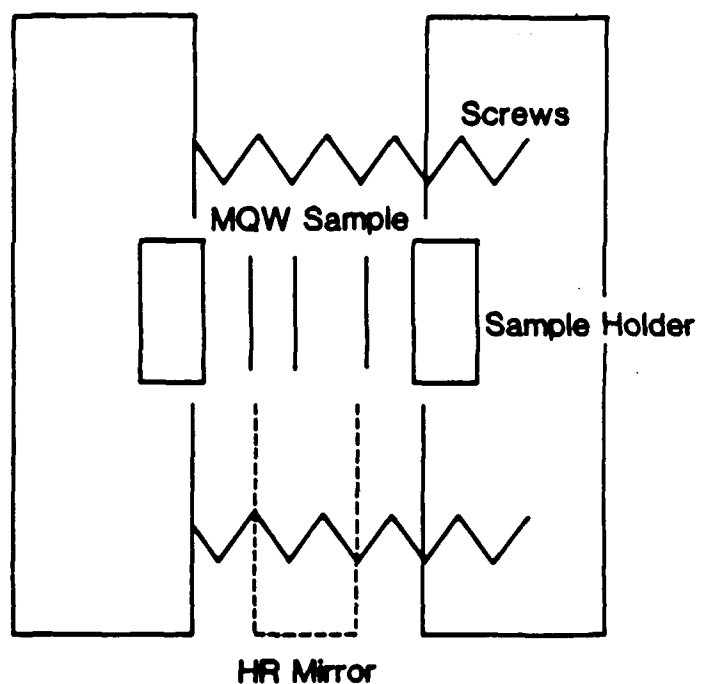


Figure 4.2 External Fabry-Perot mirrors to remove wedge nonuniformity in thickness.

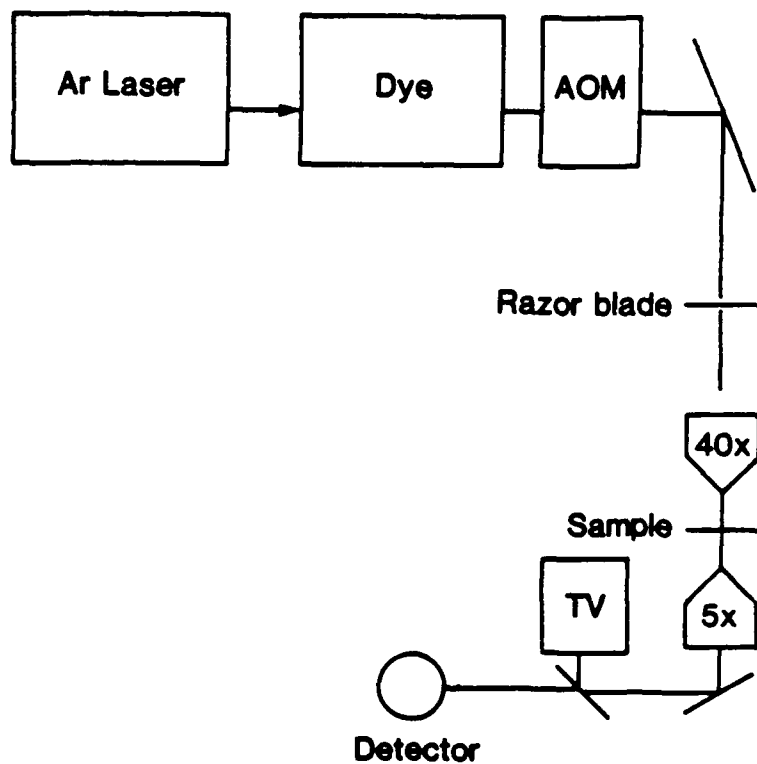


Figure 4.3 Experimental setup for testing flatness by observing shift in Fabry-Perot transmission peak.

A TRW MQW sample was used in the squeezed-sample setup. Peak wavelength versus sample position is plotted in Fig. 4.4. Further measurements show that the peak wavelength varies less than 20 Å over 800 μm. Comparisons between glued and direct-coated samples are made in Fig. 4.5 and Fig. 4.6, respectively.

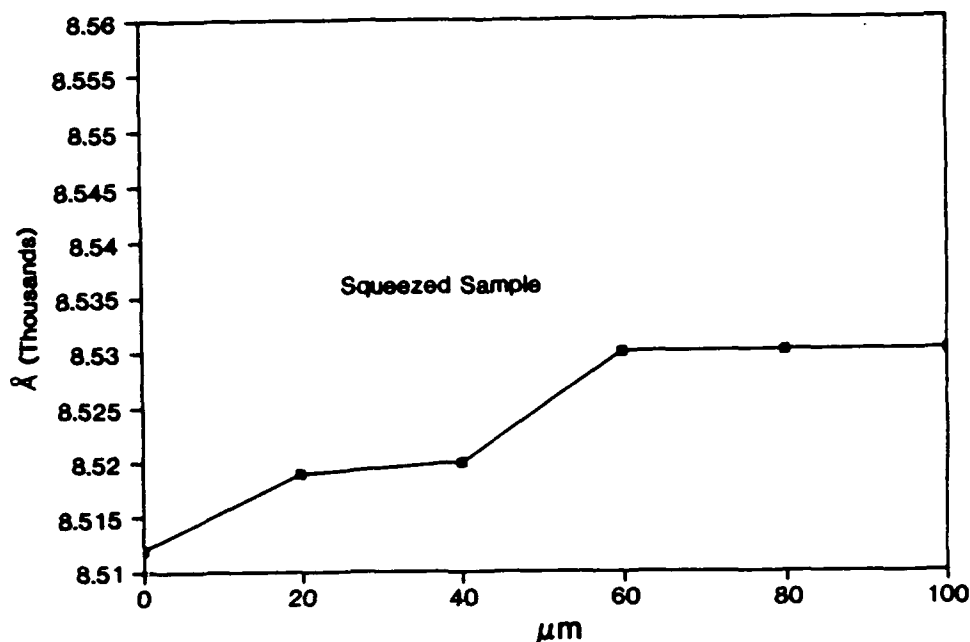


Figure 4.4 Peak-wavelength variation of etalon consisting of MQW sample with external mirrors.

4.2 Integrated Mirrors [4.2]

Direct etching of the etalon can be avoided if an integrated mirror is grown by MBE. Other laboratories have found this technique to be effective [4.3, 4.4]. Because the etching process introduces no *intracavity* thickness nonuniformity, only nonuniformity introduced by the growth process can result.

Because of their optical properties, distributed-feedback structures offer interesting possibilities as interferometric devices [4.5–4.7]. The ability to produce high reflectors with GaAs/AlGaAs quarter-wave stacks is essential for the implementation of a monolithically-grown Fabry-Perot etalon. We use the quarter-wave stack structure to grow high-reflecting mirrors on each side of a GaAs spacer layer [4.3,4.8]. In the past, Fabry-Perot etalons were constructed by sandwiching a 1-μm-thick GaAs layer between external dielectric mirrors. This procedure yielded

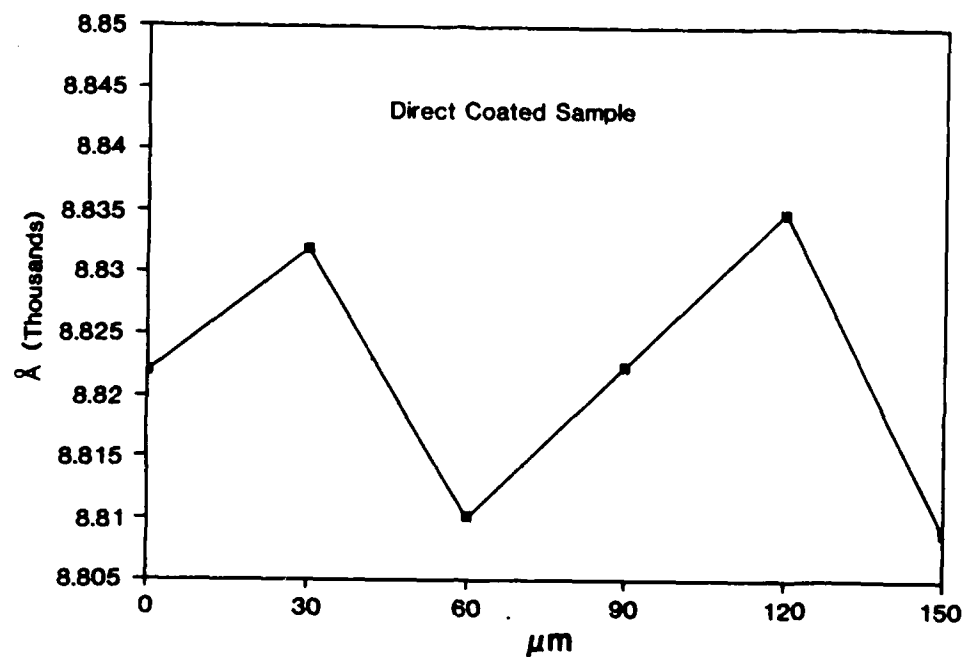


Figure 4.5 Peak-wavelength variation of etalon consisting of sample with cover-slip mirrors glued directly to sample.

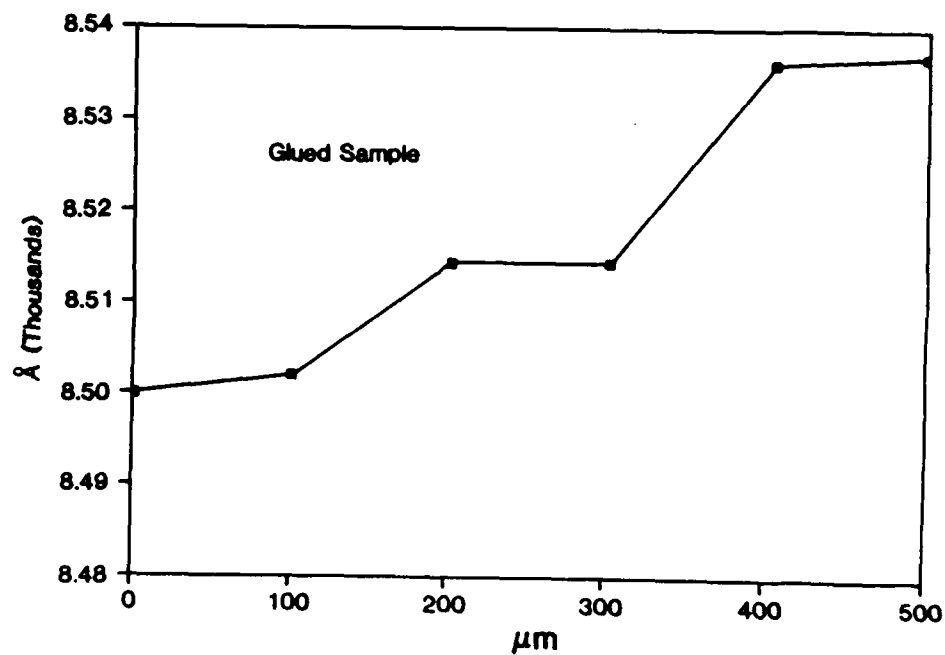


Figure 4.6 Peak-wavelength variation of etalon consisting of sample with mirrors formed by direct coating of dielectrics.

devices that were highly nonuniform; for instance, the Fabry-Perot resonance wavelength could vary by several free spectral ranges in an area as small as 50 μm . Such nonuniformity occurred either because wedges were produced in the GaAs layer during the etching process and/or because of the nonparallel alignment of the external mirrors. Although portions of such a sample might yield good nonlinear optical performance, it is difficult to imagine an array of these devices working.

The integrated Fabry-Perot concept solves the uniformity problem and produces devices comparable to the best devices produced in the past, with highly improved heat-sinking capabilities in reflection. The design of the device involves two main steps. First, one determines the reflectivity, spectral bandwidth and center wavelength of the high reflectors. Second, one chooses the spacer layer such that the resonance wavelength of the structure has the proper detuning from the optical absorption band gap to yield a high-finesse cavity with nonlinear index changes great enough to produce optical bistability.

The reflector design consists of alternating layers of GaAs and AlGaAs. If one grows materials with dissimilar indices of refraction, he can obtain optical reflections from the surfaces by virtue of this index mismatch. Ideally, one would like two materials with an index contrast as great as possible. For this reason, we choose GaAs and AlGaAs; the corresponding index mismatch is $\Delta n = 0.6$. In Fig. 4.7 we show how the reflectivity of a quarter-wave stack of GaAs and AlGaAs, with varying amounts of aluminum, changes with the number of periods in the stack. One can see that for stacks having more than 10 periods, the gain in reflectivity becomes smaller and smaller. To obtain a reflectivity of 90% over a spectral range tens of nanometers wide, at least seven periods using AlGaAs material are required. Determination of the proper detuning requires some modeling. The main criterion for the design is the achievement of high-contrast-switching operation in the reflection mode. Because the modeling program includes the nonlinear absorption and nonlinear refractive index, realistic parameters can be evaluated easily. The spacer layer itself must be a multiple of $\lambda_0/2n$, where λ_0 is the resonance wavelength and n is the refractive index of the spacer material. An intensive study of the performance of a nonlinear Fabry-Perot etalon, which includes results from the full GaAs plasma theory, has been completed [4.9].

The final design is shown in Fig. 4.8. The sample is grown entirely by MBE. During growth, the substrate is rotated in order to enhance uniformity. A short-period superlattice is grown on top of the substrate to improve the surface morphology so that subsequent layers will grow more smoothly. The success of this fabrication process relies not only on uniformity during the growth process but also

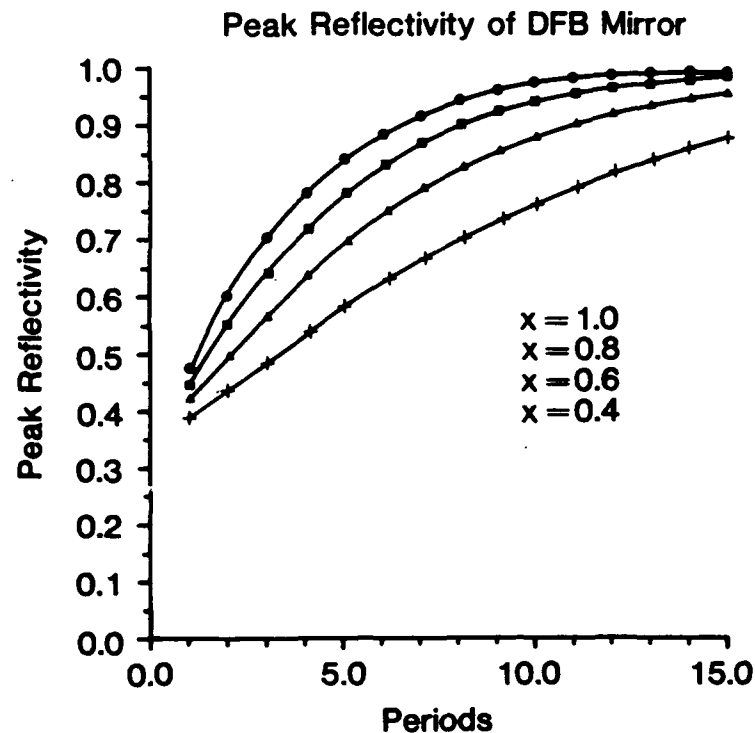


Figure 4.7 Maximum reflectivity of a GaAs-AlGaAs DFB structure as a function of the aluminum content and the number of periods in the structure.

on the accuracy of calibration of the layer thicknesses. For instance, a 5% error in calibration degrades the performance of the reflector substantially; the Fabry-Perot resonance wavelength could be spectrally shifted into the high-absorption region of the GaAs, resulting in a poor-finesse structure.

The experimental reflection spectra at different positions of a 1 cm × 1 cm section of the wafer is shown in Fig. 4.9. The theoretical performance of the structure is also indicated. As shown, the target resonance wavelength is near the designed value. More importantly, the resonance wavelength deviates less than three linewidths over the entire 1 cm² region; this is by far the most uniform performance of these devices to date. The finesse of the structure is approximately 18, which again is the highest to date in an etalon capable of optical bistability.

Integrated MBE-Grown Fabry-Perot Etalons in GaAs

GaAs	575 Å	}	x7 High Reflector
AlAs	795 Å		
GaAs	1.5975 μ		Spacer
AlAs	795 Å	}	x9 High Reflector
GaAs	575 Å		
AlAs	795 Å	}	x38
Al _{0.5} Ga _{0.5} As	100 Å		
Al _{0.1} Ga _{0.9} As	100 Å		
Al _{0.5} Ga _{0.5} As	500 Å		
Semi-insulating GaAs substrate			

Growth: Wright-Patterson AFB

Figure 4.8 Design of an integrated Fabry-Perot etalon using GaAs and AlAs quarter-wave layers with a bulk GaAs spacer layer.

4.3 Arrays With Integrated Mirrors [4.2]

We etched a 100×100 array of $5 \mu\text{m} \times 5 \mu\text{m}$ square pixels in the structure. This is done using contact photolithography and reactive ion etching. The surface of the structure is coated with a $1\text{-}\mu\text{m}$ -thick layer of a UV-sensitive photoresist. The UV light passes through a mask attached to the surface of the sample, which transfers the array pattern to the photoresist material. The photoresist is developed, leaving behind an array of pixels in photoresist. We use a BCl_3 reactive-ion-etching process to define the pixels in the GaAs structure. An SEM photograph is shown in Fig. 4.10. From the profile of a single pixel, we observe a periodic variation in contrast near the top of the pixel. This variation corresponds to the top, quarter-wave stack high reflector. The etch depth was chosen to be $3 \mu\text{m}$, so that the base surface after the etching would expose the bottom high reflector. The reflectivity of the bottom reflector is 90% over a wide spectral range, and its spectral dependence

agrees well with the predicted performance of the mirror. The experimental reflection spectrum is shown in Fig. 4.11.

4.4 Integrated Mirrors: Nonlinear Optical Performance [4.2]

The nonlinear optical performance of the device is characterized by the quality of the optical bistability and optical NOR-gate response. A typical optical-bistability trace in reflection is shown in Fig. 4.12. The dye laser is modulated by acousto-optic modulators to produce triangular pulses at a 10 kHz rate. The inverted input pulse is shown in the oscilloscope trace along with the reflected signal. The output is plotted as a function of the input to reveal the bistability loop. Bistable switching can be understood in the following qualitative way. The input wavelength is tuned to the short-wavelength side of the Fabry-Perot resonance. As the intensity in the cavity builds up, the refractive index in the spacer layer decreases. The resonance thus is shifted to shorter wavelengths, such that more light from the input enters the cavity, decreasing the index even further. This process is cumulative, and produces the sharp switching behavior in the etalon. The peak power on the etalon is 10 mW focused into a 20- μm spot. Tighter focusing of the input with higher power microscope objectives has the undesirable effect of broadening the Fabry-Perot resonance. This broadening occurs because each angle incident on the etalon sees an effectively different resonance. When the input cone of light is large (for stronger focusing), the different resonances are convolved together. The optical power requirements for nonlinear switching do not decrease with higher-power microscope objectives. However, if an optical fiber is butt-coupled to the input face of the resonator, better performance can be obtained. This improvement is possible because the light exiting the fiber is a spatially-filtered plane wave. Light reflected from the etalon is collected by the same fiber, and detected with a partially-reflecting mirror. This method can reduce the power requirements by a factor of 2 when a single-mode optical fiber with a guiding diameter of 10 μm is employed. This configuration was used to implement a fiber-etalon-fiber optical bistable fiber link (Section 9.3).

The results of an optical-NOR-gate operation are shown in Fig. 4.13. The probe wavelength is tuned to the transmission resonance of the etalon. We remove the substrate from the integrated Fabry-Perot etalon in order to perform transmission measurements. The sub-nanosecond pump pulse (88 nJ) is tuned to 817 nm for efficient carrier generation and focused to a 260- μm spot diameter. As soon as the pump is absorbed, the refractive index of the spacer decreases and the resonance is shifted to shorter wavelengths. The probe transmission decreases as the Fabry-Perot resonance shifts, because the probe wavelength is no longer resonant with the device.

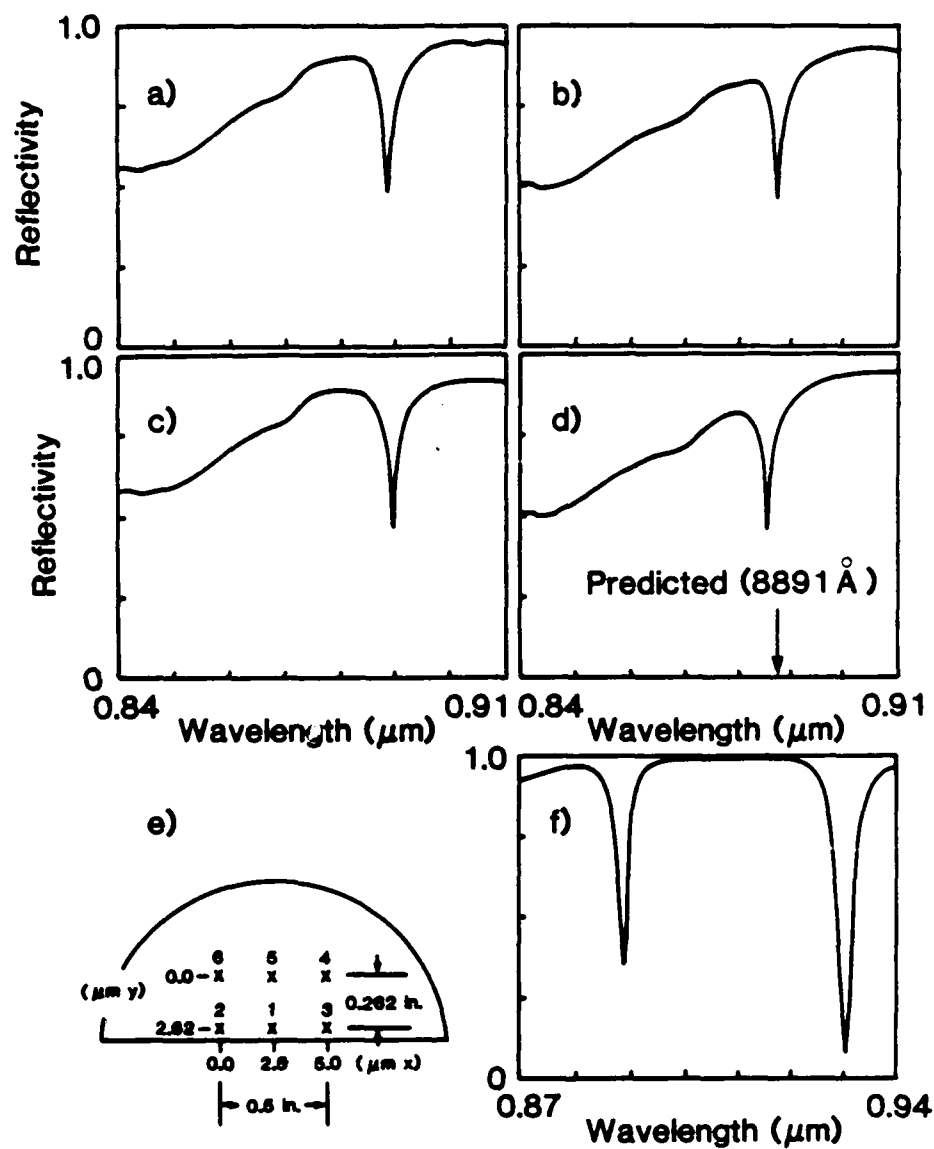


Figure 4.9 (a) - (d) Linear reflection spectra of integrated Fabry-Perot etalons taken from different portions of the same wafer. (e) Schematic representation of the wafer where the spectra in a,b,c,d, correspond to positions 1,4,3,6. (f) Theoretical linear reflection spectrum for the structure.

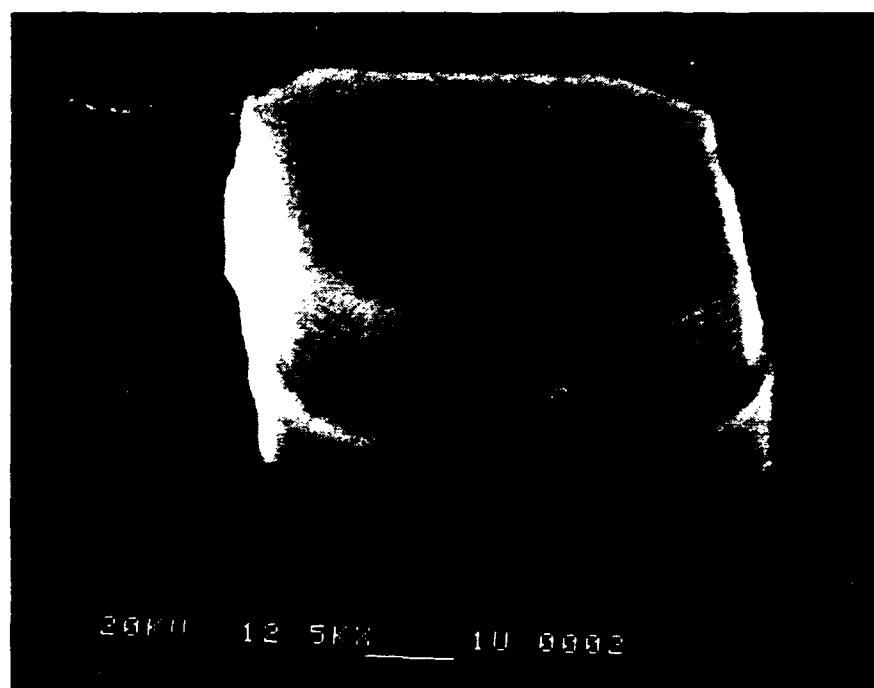
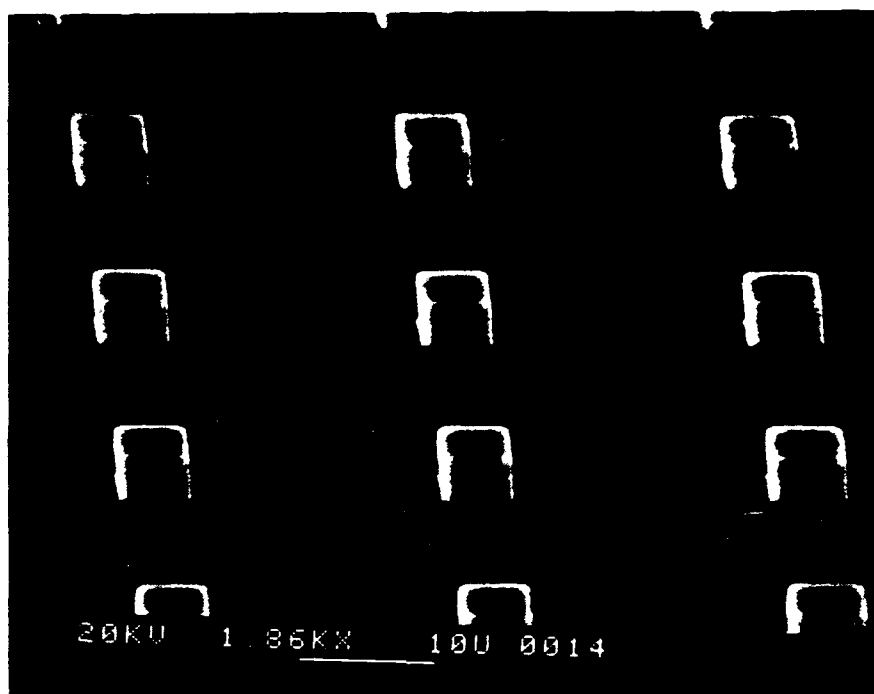


Figure 4.10 SEM photograph of an array of integrated Fabry-Perot devices. Each pixel is $5 \times 5 \mu\text{m}$ with a center-to-center separation of $20 \mu\text{m}$. The bottom picture is a magnified view of a single pixel. Note the integrated mirror at the top of the pixel.

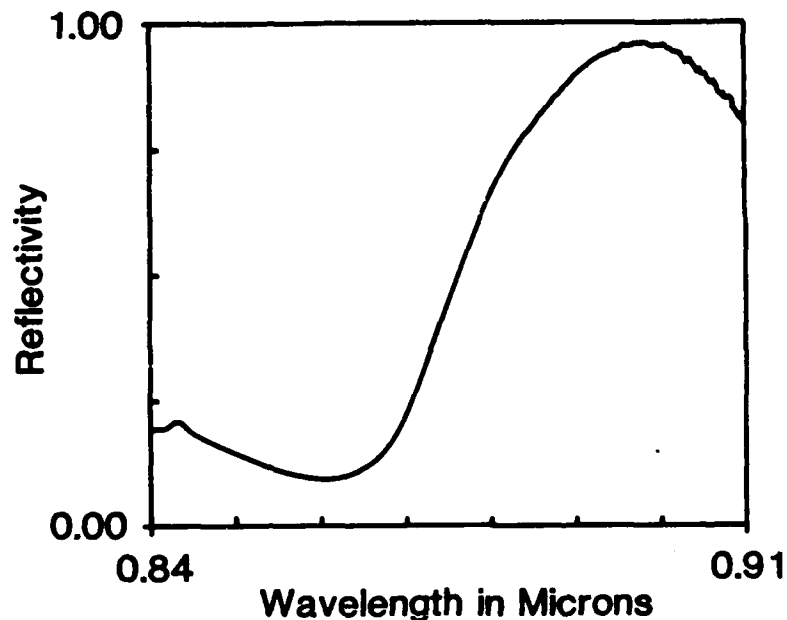


Figure 4.11 Reflection spectrum of base mirror in the integrated Fabry-Perot etalon after pixellation via reactive ion etching.

As the carriers recombine, the refractive index returns to its original value and the Fabry-Perot resonance returns to its original resonance wavelength. As the Fabry-Perot resonance returns to its original resonance wavelength, the probe transmission increases. The $1/e$ lifetime associated with the NOR-gate operation is 12 ns. Faster NOR-gate operation has been reported [4.10]. If more of the GaAs surface is exposed to nonradiative recombination, the carrier lifetime can be shortened. As was the case with the quantum wires, increasing the exposed surface leads to shorter lifetimes. We also explored this possibility in the pixellated integrated Fabry-Perot devices. In this case, only sidewall recombination contributed to the faster recombination, and yet we still observe a faster NOR-gate recovery (1 ns). The use of smaller pixels (on the order of 1 μm), should speed up the device even further.

4.5 Integrated Mirrors: Reduction in Thermal Sensitivity [4.2]

The operation of closely-spaced, bistable devices can introduce undesirable thermal effects. Thermal nonlinearities produce positive refractive index changes, which compromise the electronic nonlinearities. Unless energy can effectively be removed from the devices, these thermal effects can degrade the performance as well as produce thermal gradients in the wafer. The integrated Fabry-Perot structure

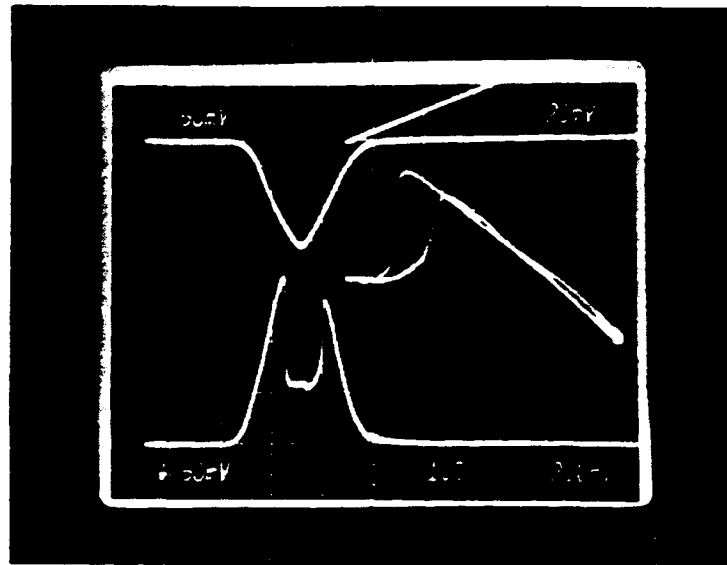


Figure 4.12 Optical bistability in reflection in the integrated Fabry-Perot etalon. The input wavelength is $\lambda = 8873\text{-}\text{\AA}$ and the peak input power is 30 mW.

operating in reflection solves some of this thermal loading problem. Heat conduction from the etalon into the GaAs substrate is efficient. This allows the observation of nonlinear optical gating with millisecond pulses.

Without any heat-sinking in the etalon, thermal effects are observed with pulses as short as 5 μs . In this monolithic etalon, we observe optical bistability with 400 μs pulses (see Fig. 4.14). Other methods of heat sinking Fabry-Perot etalons have also been reported [4.11, 4.12]. These methods, however, require additional processing steps. The integrated Fabry-Perot structure experiences a reduced thermal sensitivity simply because excess heat can be conducted into the GaAs substrate and dissipated into a much larger thermal mass.

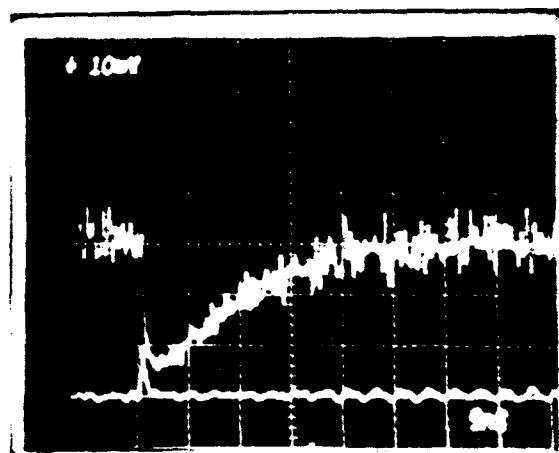


Figure 4.13 NOR-gate response of the integrated Fabry-Perot etalon in transmission. The probe wavelength is $\lambda = 888$ nm. The pump pulse is 0.5 ns full-width, half-maximum.

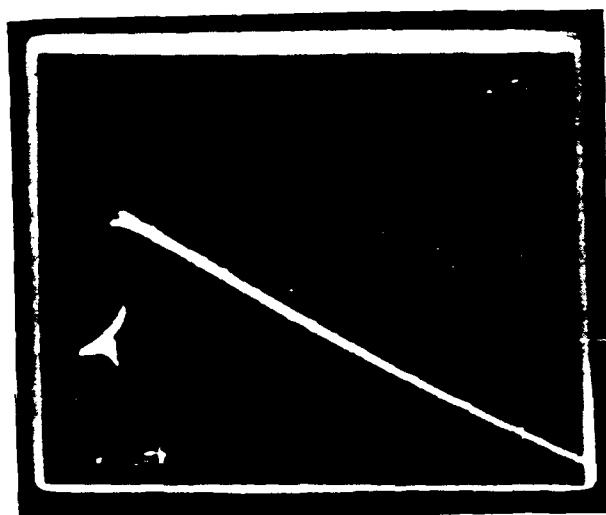
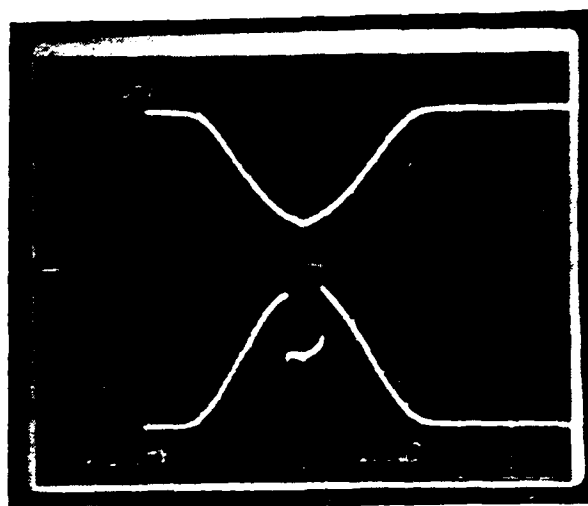


Figure 4.14 Optical bistability in reflection in the integrated Fabry-Perot etalon using 400 μ s triangular pulses. The peak input power is 40 mW and the input wavelength is $\lambda = 887$ nm. The top trace shows the time response of the switching. The bottom trace shows the narrow bistable loop.

5. CASCADING AND DIFFERENTIAL GAIN IN GaAs ETALONS USED AS LOGIC GATES

5.1 Introduction

It may be possible to use a nonlinear Fabry-Perot etalon as a switching element in a digital optical processor. These etalons may find application both in massive parallel processing and fast sequential processing [5.1]. Since the observation of optical bistability in both bulk and MQW GaAs Fabry-Perot etalons, much effort has been devoted to making these devices useful as all-optical switches. These etalons show promise because of their large nonlinearity, fast switching speeds, small size and diode-laser compatibility [5.1-5.4]. External off- and on-switching, two-wavelength gating, and fabrication of etalon and microresonator arrays have been demonstrated with these devices [5.5-5.8]. Additional requirements must be placed on such devices if a real system, or even a fairly simple circuit, is to be constructed from them. Most importantly, the etalons must exhibit cascadability, gain (or fan-out), and good contrast. These three requirements are necessary so that: (a) signal levels can be restored after each operation in order to drive the next stage; (b) output of one logic gate can be used to switch several logic gates; and (c) the signal-to-noise ratio can be kept high throughout the system.

Single-wavelength gain and cascading have been demonstrated for some materials, such as ZnS, ZnSe, InSb, allowing some simple digital optical circuits to be built (see Sections 2, and 13 through 15) [5.9-5.14]. Of course, gain also can be obtained through two-wavelength operation of GaAs devices [5.8], but then cascading is difficult because: 1) the wavelength of the output beam is different from that of the input beam in order to achieve efficient pumping; and 2) isolated resonances, which would allow the two wavelengths to be switched interchangeably, are not yet allowable.

In this section we report the experimental demonstration of single-wavelength latching circuits, with two 58-Å MQW bistable etalons operating in both the transmission (AND-gate or OR-gate) and the reflection (NAND-gate or NOR-gate) modes. The sample consisted of 180 periods of 58Å-thick GaAs and 96Å-thick AlGaAs, with a total GaAs thickness of 1.04 μm . This sample was chosen because previous measurements showed that MQWs with smaller well thicknesses exhibit greater nonlinearity (see Section 3.3). The 58Å-GaAs layer in the sample made it the narrowest-well MQW available to us. The sample exhibited the best switching

characteristics in the laboratory, despite the nonuniformity in thickness. This may be because the exciton peak in this sample is close to the wavelength at which our laser gives its maximum power. With milliwatt powers, these devices are operated at room temperature, with pulses of about one microsecond in duration. These are much shorter than those used previously for thermal bistable devices (such as ZnS and ZnSe interference filters) in demonstration experiments [5.9, 5.10, 5.14]. Even though the etalons are not optimized, and even though no special procedure is followed to stabilize the laser power, a gain of 4 is obtained. Great contrast is observed also, with the largest contrast seen $\cong 10$ for an etalon operating in transmission mode (Fig. 5.1). Thus requirements (a)-(c) discussed above are satisfied for microsecond pulses. Later in this section we discuss the use of these nonlinear etalons in large-scale parallel optical signal processing, and in fast sequential optical signal processing.

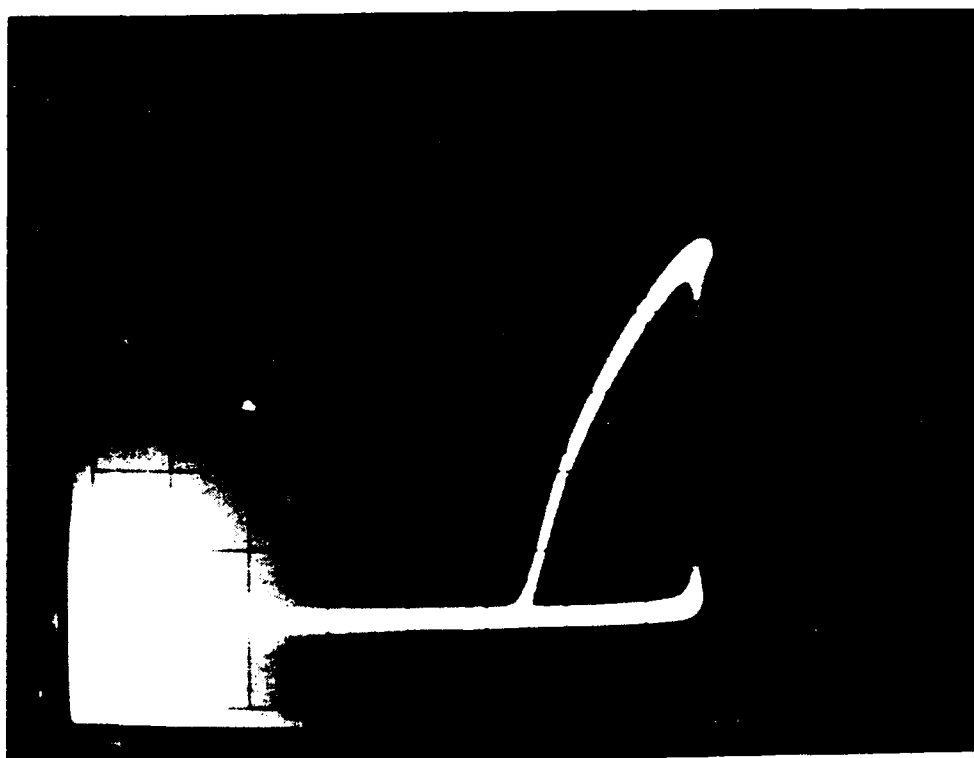


Figure 5.1 Output vs input operating characteristics of a GaAs/AlGaAs MQW nonlinear etalon showing a contrast C_d of 10. The input is a $1.5 \mu\text{s}$ triangular pulse with 45 mW peak-power.

5.2 Definitions

For a nonlinear etalon operated either in a bistable or in a thresholding mode, the gain and the contrast can be defined in terms of the power of the light pulses as follows. Suppose a weak signal beam is amplified through a nonlinear etalon by a strong holding beam; we define the device gain as

$$G_d = \frac{P_u - P_l}{P_s} \quad (5.1)$$

where P_u is the output power when the device is in the upper bistable state, P_l is the output in the lower state, and P_s is the signal-beam power (see Fig. 5.2). This definition of G_d is suitable for the steady-state operation of a device, i.e., when the input pulse length is much greater than the device response time. This gain is sometimes called *differential gain* to distinguish it from the type of gain that occurs in a laser, in which the medium itself amplifies light. Note that G_d is limited by practical considerations such as the stability of the laser amplitude and frequency, and the etalon temperature stability. Theoretically, G_d can approach infinity if one decreases P_s , but the switching time may increase because of the critical-slowness effect. Critical-slowness refers to the slower switch-on time that results when the holding power, or bias level, is near the switch-on value. In other words, for input intensities near the threshold value, the time it takes for the output to reach steady state increases. The switch-on time, then, can be decreased by employing an input intensity significantly greater than that required for switching (see, for example, pp. 221-229 of Gibbs) [5.3].

The contrast of a logic gate is defined as

$$C_d = \frac{P_u}{P_l} \quad (5.2)$$

For most applications, the contrast represents the signal-to-noise ratio of a device. In general, a large contrast is desired, especially for applications such as the nonlinear decision making in an optical associative memory (Section 14 and [5.15]). For an etalon operating in the transient mode, where the pulse length is comparable to the device response time, we can define the device gain and the contrast in a similar way, with the powers in the above expressions replaced by energies.

The minimum gain required for cascading is $G_d > 1$. Generally, for a gate to have a fan-out of N , $G_d > N$ is required if the output of the first gate is to switch

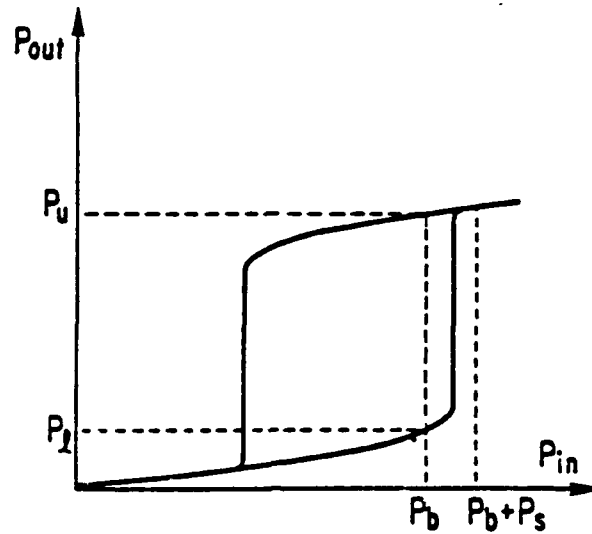


Figure 5.2 A bistable device as a latching logic gate. P_b is the bias power, P_s is the signal power, P_u and P_l are the upper- and lower-state output powers, respectively.

N succeeding gates by itself. This requirement can be reduced if the outputs of several gates are used to switch one gate (i.e., if fan-in is employed), provided the signals meet certain conditions (see Section 8.1). The contrast must also be great enough that $P_u/N > P_s$ but P_l/N is too small to switch the device. This requirement in effect determines G_d . The N required depends on the functional architecture. For example, if the simplest symbolic-substitution logic is to be implemented, $N > 2$ will be required [5.16].

GaAs nonlinear etalons can be operated in the transmission mode or in the reflection mode (Fig. 5.3). When using etalons with symmetrical cavities (like the ones used in this experiment), one can take advantage of their bidirectional operation, i.e., the transmitted and/or reflected beams from both sides may be used. Latching properties are important in this case. Latching means that the device stays in the switched state even if the small signal power is removed. Latching occurs when the device is operating in a bistable mode, with the bias beam power exceeding the switch-down power. Latching times can be controlled by setting and resetting the bias pulse, as is done in this experiment, where the device acts like a latching OR-gate in the transmission mode and a latching NOR-gate in the reflection mode.

5.3 Experimental Setup

Nonlinear Fabry-Perot etalons are made by sandwiching a 180-period 58-Å GaAs/96-Å AlGaAs MQW between two dielectric mirrors having 90%

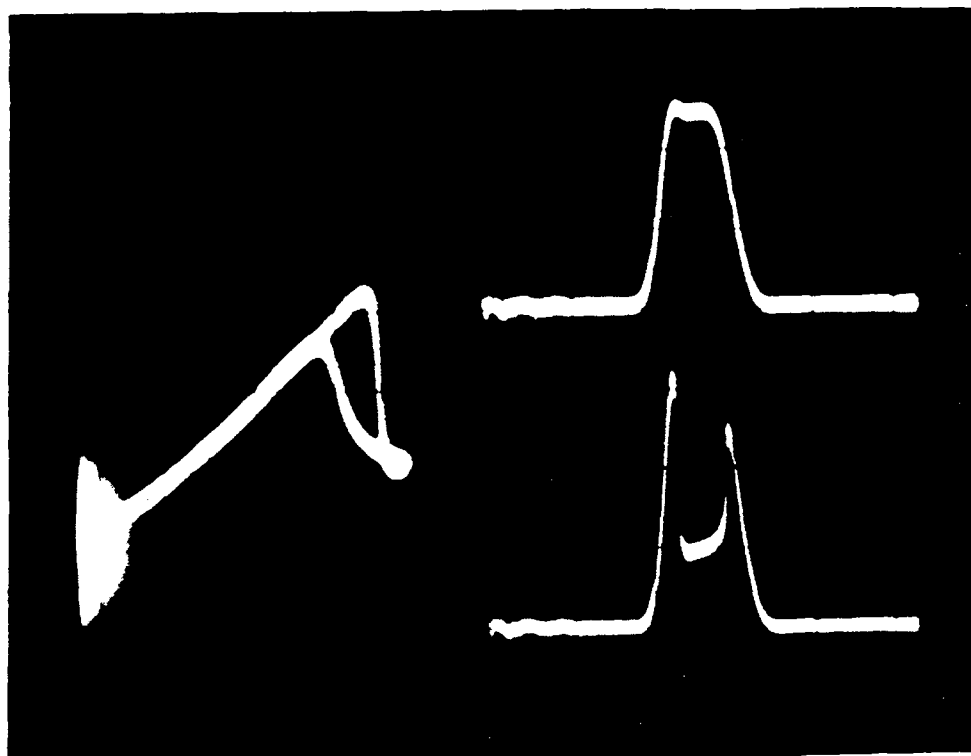


Figure 5.3 A GaAs/AlGaAs MQW bistable etalon operating in the reflection (latching NOR-gate) mode. The input pulse length is 800 ns, and the operating wavelength is 838 nm with about 35 mW input peak-power. The upper-right trace is the reflected output when the etalon is not switched, and the lower one is the output when the etalon is switched. On the left-hand side is the output vs input trace.

reflectivity. Etalons fabricated by this method usually are not uniform, but this non-uniformity makes it easier to match the cavity resonances of the two devices. Figure 5.4(a) shows the experimental setup for cascading in reflection mode. The first etalon works as a NOR-gate, and the second etalon is switched by the beam reflected from the first etalon. The input pulses are generated by a pulse generator driving an acousto-optic modulator, which diffracts a 0.4- to 4- μ s pulse from a near-infrared cw dye laser beam. This beam is split into two paths in order to switch the two etalons. If the incident laser beam is linearly polarized, beam combining is simplified. When the beam is incident perpendicular to the layers of an MQW sample, operation of the etalon is independent of polarization. This property enables one to use the polarization degree of freedom to fan-in two beams with negligible loss. The polarization degree of freedom is used to control the beam intensity. Fine

adjustment of the beam intensities is accomplished by placing a half-wave plate in front of a polarizing beam splitter. The polarization degree of freedom is used also for beam combining. The beam reflected from the first etalon is coupled into the second etalon by rotating its polarization by 90° , using a quarter-wave plate and a pair of polarizing beam splitters. After the beam passes the quarter-wave plate twice, it is combined with a holding beam of opposite polarization, which switches the second etalon. The laser beams are focused to spot sizes of about $15\text{ }\mu\text{m}$ in diameter with $10\times$ microscope objectives; the spots are monitored by means of a TV camera. The input power to the first stage is varied by rotating the half-wave plate or by changing the aperture size of the pinhole so that the etalon can be switched on or off. The change of the output signal from the first stage then is measured and compared with the change of the output signal of the second stage to determine the gain. Similar setups are used for the cascading experiments, with both etalons working in transmission mode as OR-gates, where the transmitted signal of the first etalon is used to switch the second etalon, as shown in Fig. 5.4(b).

5.4 Discussion of Experimental Results

Optical bistabilities are observed in the etalons described above, with microsecond pulses in the 840-nm-wavelength region; the minimum input power for bistability is about 6 mW. The shortest pulse used to observe optical bistability is 400 ns, and is limited by the pulse generator. A typical result of the cascading experiment is shown in the oscilloscope traces of Fig. 5.5. Note that these results are for cases where the devices are operated in the bistable, or latching, mode. It is possible to adjust the experimental parameters (e.g., the etalon detuning) to make the bistable loop extremely narrow, to approach the case of thresholding, or non-latching, operation. We have demonstrated thresholding operation experimentally. The bias-beam powers in these experiments are 20-40 mW to yield sufficient gain, and the etalon with the lower threshold is used in the first stage. The etalons were formed by sandwiching, yielding etalons of uniform thickness. The spot with best switching was arrived at by trial and error. Later, we made more uniform etalons using the all-integrated, MBE-grown mirrors (see Section 4.2). Gains greater than 2 are obtained in both cases, as demonstrated in Figures 5.4(a) and 5.4(b), but the contrast is less in the reflection mode because of the low finesse of the cavity. The largest gain observed is 4, where a 0.25 mW change of the output signal of the first stage induced a change of 1 mW in the output signal of the second stage. The contrast at these power levels is 5-8; contrast can be increased to about 10 through the use of greater input powers (45-50 mW). Latching times usually are set on the order of

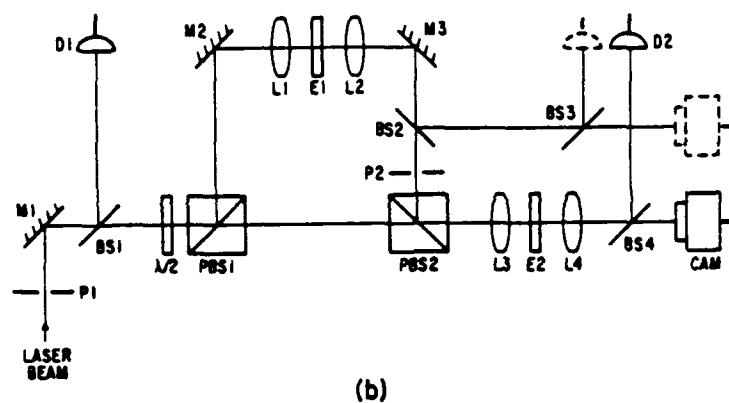
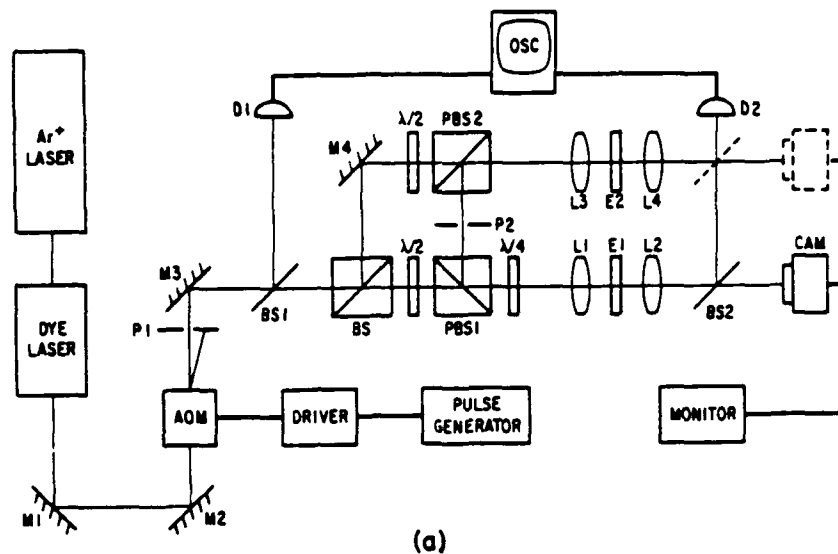


Figure 5.4 (a) The experimental scheme for cascading in reflection mode. *M*: mirror; *P*: pinhole; AOM: acousto-optic modulator; BS: beam splitter; $\lambda/2$: half-wave plate; $\lambda/4$: quarter-wave plate; PBS: polarizing beam splitter; *E*: nonlinear etalon; *D*: detector; CAM: camera; OSC: oscilloscope. (b) The setup for cascading in transmission mode.

microseconds to keep the device thermally stable. Similar results are expected for etalons made from bulk GaAs material.

We found that focusing the output beam from the first device onto the second device is crucial for experimental observation of gain. Unlike the single-beam or two-beam experiments, where both bias and signal beams come directly from the laser source, in this experiment the signal beam for the second stage is the output from the first stage, reshaped spatially and temporally by the first-stage nonlinear etalon. Spatial reshaping is attributed to the Gaussian profile of the input beam and

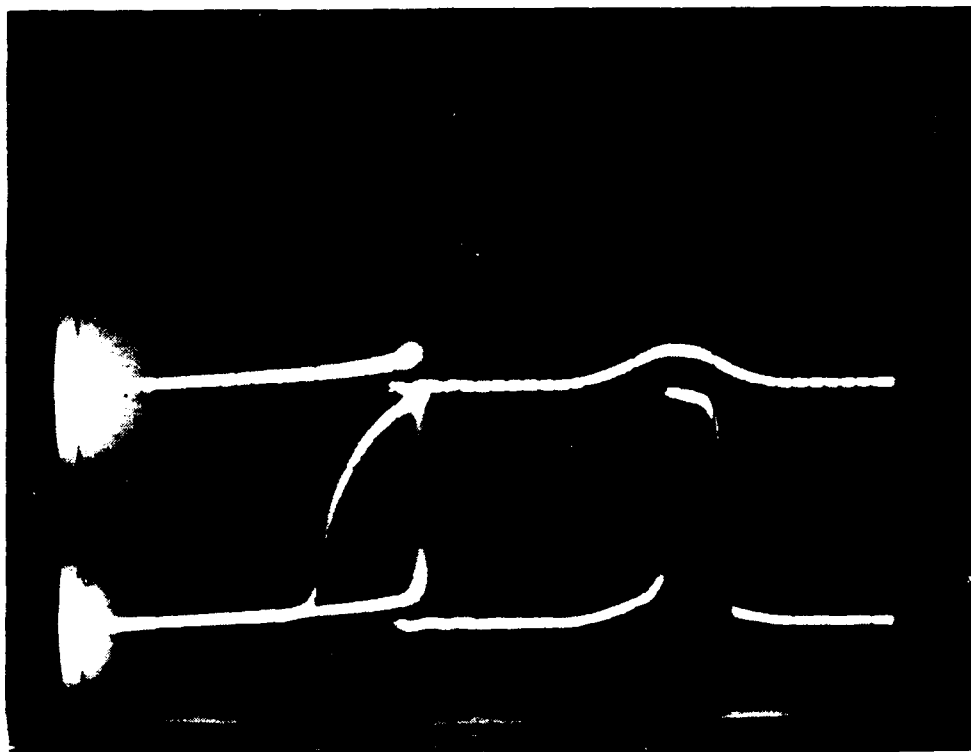


Figure 5.5 Results of the cascading experiment corresponding to Fig. 5.4(b). The traces on the left are plots of output vs input, and the traces on the right are outputs from the second stage. The upper traces represent the case when the first stage is "off," and the lower traces represent the case when the first stage is "on." The total input peak power for each device is about 30 mW, and the input pulse length is 1.5 μ s.

to the nonlinearity of the etalon. Ideally, the limiting characteristics of the bistable device flatten the Gaussian profile of the beam after it passes the first etalon, making the signal beam on the second etalon more uniform. Experimental observations, however, show that the signal beam on the second etalon is not always spatially uniform after passing through the various optical components. This non-uniformity does not seem to inhibit switching of the etalons greatly, perhaps because the beams are focused tightly and because carrier diffusion is rapid, making the profile of the much-stronger bias beam the dominant factor. This would explain why the best result is always obtained when the signal beam is focused at the center of the bias beam, in which case the signal beam has the greatest influence on the bias beam. The temporal pulse-resaping by the first-stage etalon makes the rising edge of the signal beam much steeper than that of the bias beam, because the etalon switch-on

time is much shorter than the rising edge of the input pulse. For a given pulse area, therefore, this reshaping may even be desirable for switching on the next device (for the quasi-steady-state case).

At room temperature, regenerative pulsations are observed for pulses longer than a few microseconds. This means that without the use of heat sinking, cw operation in the "real" sense, i.e., infinite latching time, is difficult to achieve. Yet, latching times in the microsecond range may be considered "quasi-cw," at least for fast sequential processing applications, such as compare-and-exchange type operations (Section 15 and [5.17]), with picosecond pulses. For example, if mode-locked pulses of 10-ps are used with a latching time of 1 μ s for the bistable gates, then $1\mu\text{s}/10\text{ps/pulse} = 10^5$ pulses. Even if one assumes a duty cycle of 1:100, i.e., pulses are 1000 ps apart, one thousand pulses can be routed before the thermal effect comes into play (power dissipated by the pulses is roughly 5 mW if each pulse has a power of 10 pJ). Reducing the pulse spacing to 100 ps would increase the data rate to 10 Gbit/sec. Much longer latching times can be achieved through the use of heat sinking.

If no special procedures are followed to stabilize the dye-laser output, the maximum gain that can be measured is limited by the noise in the laser output. A detailed analysis of the influence of noise on transmission characteristics requires a statistical treatment. In order to prevent noise-induced switching in practice, it is sufficient to keep the bias-beam power $P_b < P_{su} - P_n$, where P_{su} is the switch-up power and P_n is the peak power of the noise. Therefore, the noise level defines a minimum value for the power P_s of the switching beam and, hence, a maximum value for the observable gain. Let T_u and T_l be the etalon transmission at upper- and lower-state, respectively, then the device gain given by Eq. (5.1) can be written as

$$G_d = \frac{(T_u - T_l)P_b}{P_s} . \quad (5.3)$$

In the presence of noise, one must have

$$P_s > P_n = \gamma P_b , \quad (5.4)$$

where $\gamma = P_n/P_b$, to ensure reliable operation. Thus the largest device gain (G_{\max}) observable in the presence of noise is, combining (5.3) and (5.4),

$$G_{\max} < \frac{T_u - T_l}{\gamma} . \quad (5.5)$$

We can see from Eq. (5.5) that G_{\max} is determined by (a) the noise in the laser beam, and (b) the difference between the upper- and lower-state transmission of the etalon. As an estimate, if the laser noise level is 1%, then for an etalon with 10% upper-state nonlinear transmission (T_u) operating with a contrast of about 5 near the switching threshold, the maximum observable gain is about 8. We observed fluctuations of 1%-3% in our dye laser, depending on how precisely the laser was aligned. When powering the etalons with diode lasers, the laser must be made stable. The stability of the laser depends on the stability of the current source and on its temperature stability. When these parameters are controlled, the lasers can be operated with less than 1% fluctuation. Diode lasers are much more sensitive to feedback than are dye lasers, however; great care must be taken (e.g., through the use of a Faraday isolator) to avoid feedback if the stabilization circuit is to be successful.

All etalons are operated in the bistable (latching) mode in this experiment, although cascading can be accomplished by operating the etalons in the thresholding (non-latching) mode. In practice, bistability may be preferred for several reasons. Bistability serves as a natural memory, which greatly simplifies the circuitry where latching is required. Bistability also can suppress input noise through its contrast and hysteresis, or can increase the gain by means of its large slope. In some senses, bistability corresponds to the Schmitt trigger used in electronics. Effective thresholding, however, will be required when the etalon is operated within picosecond time scales.

5.5 Possible Applications

Several applications may be envisioned, each placing different requirements on the bistable etalons. For example, symbolic-substitution logic (Section 13 and [5.9, 5.10, 5.16]) requires the simultaneous operation of large arrays (probably $> 100 \times 100$) of pixels. Good uniformity over a relatively large etalon area (on the order of 1 cm^2) is required. One may want to reduce the speed requirement for a single gate, in exchange for a lower operating power, as the pixels are densely packed. High throughput may be achieved in this scheme through the use of simple interconnection patterns and low fan-out requirements. Similar applications include spatial-light-modulators with sub-microsecond (single-wavelength, cascable) or even picosecond (two-wavelength) addressing times. Another application is high speed sequential

processing, which takes advantage of the fast switching speed of a single device. An example is the all-optical compare-and-exchange switch (Section 13 and [5.17]). With GaAs optical logic gates, the data rate of a single self-routing channel can be 1 to 10 Gbit/sec (assuming 1000-100 ps pulse spacing, as discussed in Section 5.4). The problem of picosecond cascading of passive nonlinear devices is not yet solved (see Section 4). It is conceivable, however, to make use of their fast switching speeds in logic decisions (such as gating and wavelength conversion), and then to amplify the output signals at the same rate to maintain the overall bandwidth of the system. The development of high-speed and high-gain laser amplifiers suggests a means to overcome the pulse energy loss [5.18]. Thus optical switches may find some applications in fast sequential processing.

5.6 Conclusion

In conclusion, we have demonstrated room-temperature single-wavelength optical latching circuits using GaAs MQW etalons. The wavelength and the switching power used for cascading the devices imply that diode lasers can be used as light sources, and small-scale compact demonstration digital optical circuitry can be built to operate much faster than optical circuitry built previously with thermal devices. With picosecond switch-on times, GaAs latching gates can be used in combination with other devices operating in a non-latching, high-speed mode, to provide the high data rates (\geq GHz) required for future high-bandwidth photonic switching systems.

6. THEORETICAL ANALYSIS OF DIFFERENTIAL GAIN IN GaAs ETALONS AS A FUNCTION OF PULSE LENGTH

To study the variation of the differential energy gain as a function of the temporal pulse width, numerical simulations of all-optical room-temperature GaAs devices were performed. In these simulations the transmission of a GaAs Fabry-Perot resonator of mirror reflectivity R and length L is computed as

$$I_t = \frac{I_0(t)(1-R)^2}{(e^{\alpha(\omega,N)L/2} - R e^{-\alpha(\omega,N)L/2})^2 + 4R \sin^2(\delta + \omega\Delta n(\omega,N)L/c)} \quad (6.1)$$

where δ is the linear phase shift (detuning) and $I_0(t)$ is the input intensity. To obtain Eq. (6.1), the intra-cavity light-field dynamics are adiabatically eliminated. For this approximation to be valid, it is restricted to the case where all other variables change slowly within the time it takes for a round trip through the cavity. For a study of dynamic effects on the round trip time scale, one must replace Eq. (6.1) with the dynamic equations for the forward and backward travelling field inside the cavity [6.1].

Through the semiconductor absorption coefficient $\alpha(\omega,N)$ and the nonlinear portion of the refractive index $\Delta n(\omega,N)$, the transmitted light intensity is coupled to the electron-hole-pair density N in the medium. For the present study we use the plasma theory [6.2] to compute the optical nonlinearities of bulk GaAs at room temperature. The results have been shown to explain the experimental data [6.3] (Section 3).

To relate the excitation density to the light intensity I inside the resonator, we use the simple rate equation

$$\frac{dN}{dt} = -\frac{N}{\tau} + \frac{\alpha(\omega,N)}{\hbar\omega} I, \quad (6.2)$$

where τ is the carrier relaxation time. Equations (6.1) and (6.2) are solved for both pulsed and steady-state excitations. The numerical results for the transmitted intensity as a function of the input intensity are plotted in Figs. 6.1(a)-(c) for input pulses at a frequency ω well below the exciton resonance and with pulse widths of 10τ , 100τ , and 1000τ , respectively. Results for the corresponding cw characteristics are shown in Fig. 6.1(d). Curves 1 through 4 in each figure are hysteresis curves

obtained for different detunings of the etalon eigenfrequency, ω_R , with respect to the excitation frequency, ω . The detuning is defined as $\Delta\omega = \omega_R - \omega$, where ω has been fixed for computational convenience and the resonator eigenfrequency has been varied. The curves in Fig. 6.1 reveal a sequence one would see experimentally if the laser excitation spot were moved across a slightly wedged sample. Curves 2 and 3 display well-developed bistable loops similar to those observed in experiments under the assumed operating conditions. These loops become wider for shorter pulses because of dynamical hysteresis, and they show the interplay between the dispersive tuning of the resonator and the saturation of the absorbing medium. In curve 2, the initial detuning of the resonator is such that the peak transmission of the Fabry-Perot resonator nearly coincides with the excitation frequency at the same intensity I that saturates the semiconductor absorption. This mixed dispersive and absorptive behavior leads to a much higher transmission for the device, but at the price of somewhat increased switch-on power. As shown in curves 1 and 4, the dynamical hysteresis effect can produce seemingly bistable behavior that vanishes when longer pulses are used. Qualitatively, these results are in agreement with those of Ref. [6.1], where a different medium nonlinearity has been analyzed.

To obtain the differential energy gain for pulsed operation, we solve Eqs. (6.1) and (6.2) assuming two-pulse excitation. One pulse, the so-called bias pulse, is used to bring the device near the switching threshold, and the second pulse acts as the switch pulse. The differential energy gain then is defined as the total energy transmitted in the presence of both bias and switch pulses minus the energy transmitted in the presence of the bias pulse alone, divided by the switch-pulse energy. This definition is a reasonable extension of the steady-state gain of Eq. 5.3 to the pulsed regime. Examples of the results for Gaussian-shaped pulses are plotted in Fig. 6.2. The curves show that as pulse length is shortened, differential energy gain decreases. For example, if the carrier lifetime is 10 ns, curve 1 is for a 1 μ s pulse, giving a maximum gain of 3.7; reducing the pulse length to 100 ns eliminates the gain. In principle, the problem could be solved by reducing τ to 1 ps and using 100 ps pulses. Unfortunately, the power would then scale inversely with τ , that is, bistability would require ≈ 100 W focused to a few microns.

In later simulations we used switching pulses much shorter than the carrier lifetime and longer bias pulses to investigate the possibility of increasing the differential energy gain by regulating the arrival of the switching pulse relative to the bias pulse. The increase was negligible; the differential energy gain never reached the value 2. These studies indicate that passive nonlinear etalon devices cannot be expected to exhibit useful differential energy gain for operating pulses

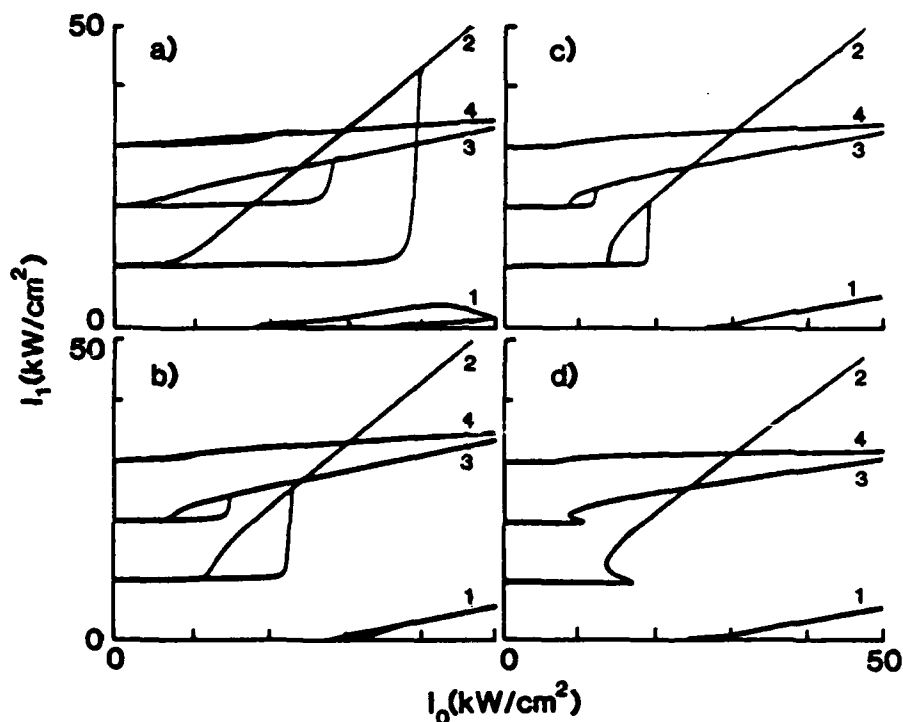


Figure 6.1 Transmitted intensity versus input intensity computed for a $\approx 2\mu\text{m}$ GaAs etalon at room temperature for the excitation energy $\omega = 883.7\text{ nm}$ (1.4032 eV) well below the exciton resonance at 873.2 nm (1.420 eV). Figs. 6.1(a)-(c) are obtained assuming pulsed excitation with a triangular pulse of full width 10τ (a), 100τ (b), and 1000τ (c). Figure 6.1(d) shows the steady state results. Curves 1 - 4 in each figure are for the detunings $\Delta\omega = \omega_R - \omega = -10.7\text{ nm}$ (-0.0170 eV), -8.95 nm (-0.0142 eV), -7.25 nm (-0.0115 eV), and 5.0 nm (-0.008 eV), respectively. ω_R is the resonator eigenfrequency and the mirror reflectivity $R = 0.9$. The baseline for the transmitted intensity in curves 2, 3, and 4 has been shifted by 10, 20, 30 kW/cm^2 , respectively.

shorter than 10τ . In order to achieve single-wavelength, cascable, optical logic gates that can operate with picosecond pulses, it will be necessary to dramatically reduce the carrier lifetimes in the semiconductor material.

Because the bistable behavior of GaAs MQW etalons is similar to that of bulk GaAs etalons, one can predict that the above conclusions will hold, at least qualitatively, for the GaAs MQW etalons used in the experiment. Indeed, we observed a decrease in the differential energy gain as the pulse length was shortened to several hundred nanoseconds. Efforts to reduce the carrier lifetime already have

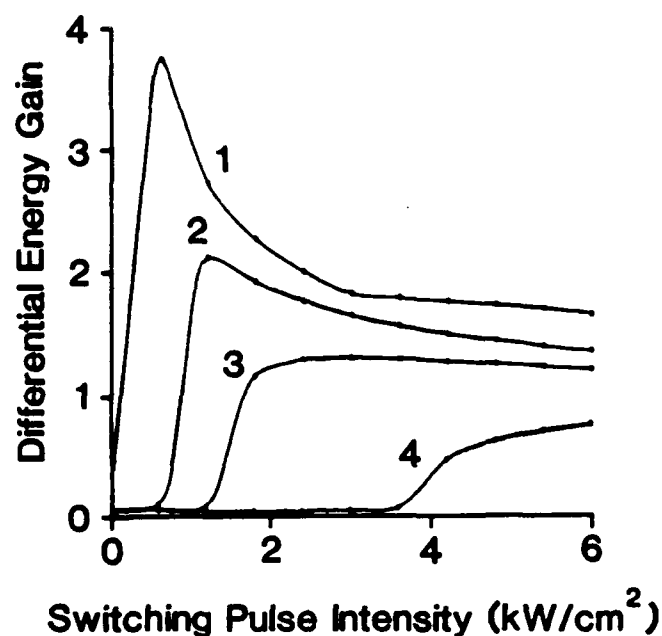


Figure 6.2 Differential energy gain versus switch-pulse intensity for different ratios of temporal pulse width Δt to carrier relaxation time τ . The parameters are those of curves 2 in Fig. 6.1 and the peak intensity of the bias pulse was kept at 17.1 kW/cm^2 , i.e., just below the steady state switch-up value. (1) $\Delta t/\tau = 100$; (2) $\Delta t/\tau = 50$; (3) $\Delta t/\tau = 25$; (4) $\Delta t/\tau = 10$.

been made by reducing the surface-to-volume ratio of the semiconductor material, thus providing more sites for non-radiative carrier recombination. Examples are the use of thin samples without protective AlGaAs "windows" [6.4] and increasing surface recombination rates through the etching of surface features [6.5]. The reduction of pixel size will also serve this purpose [6.6]. Another promising approach could be the use of semiconductor micro-crystallites. For the case of II-VI compounds it has been shown that the lifetime is reduced by two to three orders of magnitude when reducing the crystal from bulk to micro-crystallites of the order of around 100 \AA [6.7].

7. ENHANCEMENT OF DYNAMIC DIFFERENTIAL GAIN OF GaAs ETALONS BY ANGLE TUNING OF THE SWITCH BEAM

7.1 Introduction

The dependence of differential gain in GaAs bistable etalons on the incident angle of the switching beam is studied theoretically for the case of pulsed, single-wavelength operation. Both dispersive and absorptive aspects of the etalon are included. Simulations predict that differential-gain characteristics can be improved significantly over the normal-incidence case by utilizing angular separation of the pump and switching beams. The limitations of the plane-wave approximation and the potential effect of beam walk-off are discussed.

The feasibility of GaAs Fabry-Perot etalons for use in all-optical digital computing depends strongly on the signal gain and switching speed attainable in such devices. In Section 5, we reported the observation of signal power gains of up to 4 in room-temperature GaAs etalons using quasi-steady-state microsecond pulses of a single wavelength (Section 5 and [7.1]). In Section 6, we concluded that for pulsed operation, the dynamic effects are always relevant in describing the device response attributed to critical slowing down and to finite material response time. In addition we predicted, on the basis of numerical simulations, that the gain observed would vanish as the pulse duration decreased towards the carrier lifetime of the material. Carrier lifetimes must be reduced significantly, we concluded, if such devices are to exhibit useful differential energy gain for picosecond pulses.

The switch-beam angle tuning analyzed in this section does not circumvent that conclusion, but it is a way to increase the gain at any given ratio of pulse length to carrier lifetime. In this approach, the incident angle of the switching beam is varied with respect to that of the pump, or "holding," beam. This variation essentially reproduces the resonant characteristics of two-wavelength operation, in which the effect of the switching beam on the etalon detuning is maximized, without sacrificing cascability. The effect of angle tuning on the steady-state switching characteristics of etalons has been explored (see References [7.2-7.4]), where a Kerr-type model of the material nonlinearities is used.

We performed dynamical simulations of incident angle-dependent switching behavior in bistable GaAs etalons, using the microscopic plasma theory [7.5] to model the carrier-density-dependent changes in both the absorptive and dispersive characteristics of the cavity. For devices based on band-edge nonlinearities (such as

in GaAs), the light-matter dynamics can differ significantly from those found in a Kerr medium, because of saturation of the absorption and because of the nonlinear index change. Consideration of these effects allows a more realistic appraisal of operational tradeoffs critical to device performance.

We begin by discussing modifications that must be made to the standard Fabry-Perot formula to account for the effect that varying the incident angle has on the absorptive and dispersive characteristics of the etalon. We then demonstrate the behavior of the gain as a function of angle, and show how operation at an optimal angle reduces the required switching intensity, making increased gain possible. We conclude by discussing the possible implications of these results for future work in this field.

7.2 Including Angle Dependence in the Fabry-Perot Formula

For a single beam incident normally at the surface of a Fabry-Perot etalon of length L , the transmitted intensity is given by Eq. (6.1) [7.6]. Here $\alpha(\omega, N)$ and $\Delta n(\omega, N)$ are the absorption and the change in the refractive index, respectively, of the material filling the Fabry-Perot cavity; ω is the incident laser frequency; N is the carrier density; δ is the detuning; and R is the reflection coefficient for the etalon. $I_0(t)$ is the incident beam intensity. We assume a Gaussian time dependence for $I_0(t)$, with the temporal full-width at half-maximum (FWHM) equal to t_0 .

The cavity length L enters into the dispersion and absorption terms in the denominator of Eq. (6.1) as a result of the two assumptions presented below. We take the angular dependence to alter the effective etalon length L alone (because the angles considered here will in any case be small, we can neglect the variation of R and T with θ). It is therefore critical to decide how L must be modified in each case.

a) Net phase shift between sequential passes through the cavity

The L in the phase-shift (dispersion) term arises out of the calculation of the optical path length difference $2L$ between a given wave and its predecessor which already has passed through the cavity. For non-normal incidence, this path difference is easily calculated [7.7] to be $2L\cos(\theta)$, where θ is the internal angle between the propagation direction and the normal axis through the etalon. Thus we must multiply L by $\cos(\theta)$ in the dispersion term.

b) Net absorptive loss in traversing the etalon

The usual boundary conditions for the cavity relate the field at the output face of the etalon to the field incident on the input face through the standard Beer's law intensity losses encountered in travelling a distance L from $z = 0$ to $z = L$ in one

pass. For non-normal incidence, the wave will travel a greater distance before being reflected at the output face, and in subsequent passes through the cavity. The effective length over which absorption occurs will increase to $L/\cos(\theta)$. Thus we must divide L by $\cos(\theta)$ in the absorption term:

$$E_F(z=L) = E_F(z=0) e^{i\beta/2} e^{-\alpha L/2\cos(\theta)} \quad (7.1)$$

With these modifications included, the formula for the transmitted intensity becomes: [7.8]

$$I_t(t) = \frac{I_0(t)T^2}{(e^{\alpha(\omega,N)L/2\cos(\theta)} - R e^{\alpha(\omega,N)L/2\cos(\theta)})^2 + 4R\sin^2(\omega L\cos(\theta)(n_0 + \Delta n(\omega,N))/c)} \quad (7.2)$$

7.3. Differential Energy Gain as a Function of Internal Angle Between Beams

To study the behavior of the signal gain as the internal angle is varied, we assume the geometry of Fig. 7.1. The pump beam is incident normally on the sample, while the switching beam strikes the etalon at an angle θ' , entering the material at angle θ with respect to the pump beam, where θ and θ' are related by Snell's law. For the field-matter interaction within the cavity, we solve Eq. (7.3) separately for each beam, taking $\theta = 0$ for the pump beam. These equations are coupled with the spatially averaged rate equation for the carrier density:

$$\frac{dN}{dt} = -\frac{N}{\tau} + \frac{\alpha(\omega,N)}{\hbar\omega} I_{tot} \quad (7.3)$$

where τ is the carrier lifetime in the material and I_{tot} is taken as the sum of the pump and switching intensities inside the resonator:

$$I_{tot} = (I_{u(pump)} + I_{u(switch)}) \frac{1+R}{T} \quad (7.4)$$

When solved in the full dynamic case, Eqs. (7.2), (7.3), and (7.4) yield the transmitted pump and switching intensities as a function of time. Although both beams contribute to the total intensity within the cavity, the transmitted switching pulse emerges from the sample at an angle θ' and, presumably, will not participate usefully in optical logic operations at the next circuit element. Thus, only the transmitted pump beam is used in determining the gain.

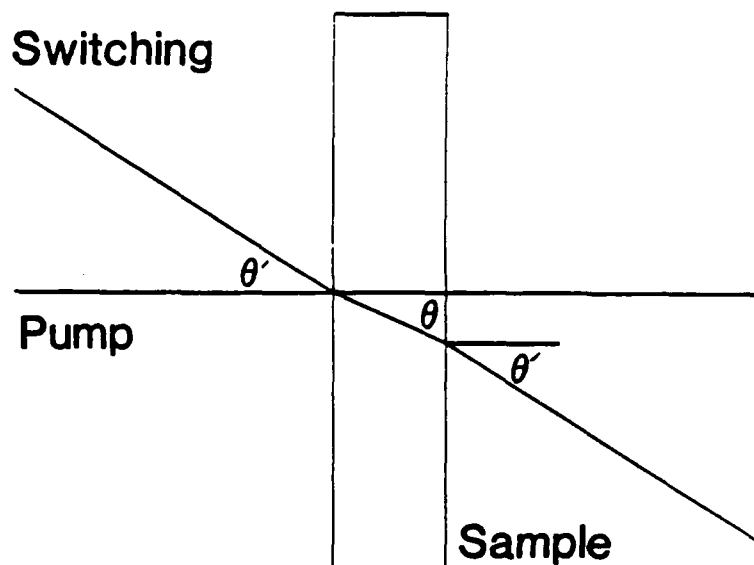


Figure 7.1 Etalon geometry for non-normally incident switching beam.

The gain calculated in this case is differential energy gain. If we define $E_t(b+s)$ as the energy in the transmitted pump pulse when a switching pulse of energy E_s is applied, and if we define $E_t(b)$ as the energy in the pump pulse transmitted when no switching pulse is used, then the differential gain is simply $G_d = [E_t(b+s) - E_t(b)]/E_s$, as in Section 6. In Fig. 7.2, we plot this value versus the internal angle θ between the beams. The parameters used (frequency, etalon length, bias and switch pulse intensities) were chosen to optimize the gain for the case $\theta = 0^\circ$. A clear maximum in the curve occurs at $\theta = 5^\circ$, where G_d has increased from the normal-incidence value of 3.8 to more than 5.0.

The origin of the gain enhancement in this case can be found in Eq. (7.2). For the switching pulse, the $\cos(\theta)$ factor in the dispersion term can be viewed as an effective downshift in the frequency ω . Assuming a laser wavelength of 883.6 nm, an internal angle of 5° between beams will result in an effective increase of the wavelength to $883.6/\cos(5^\circ) = 886.9$ nm. In the switching process we increase the carrier density from 10^{15} cm^{-3} to approximately $1.2 \times 10^{18} \text{ cm}^{-3}$, shifting the Fabry-Perot peak from around 892.9 nm to about 886.5 nm, as shown in Fig. 7.3. Angle-tuning of the switching pulse allows it to act on the etalon in a frequency region where transmission is far greater than for the normal-incidence case. Thus $E_t(b+s)$ is increased and the gain is correspondingly enhanced.

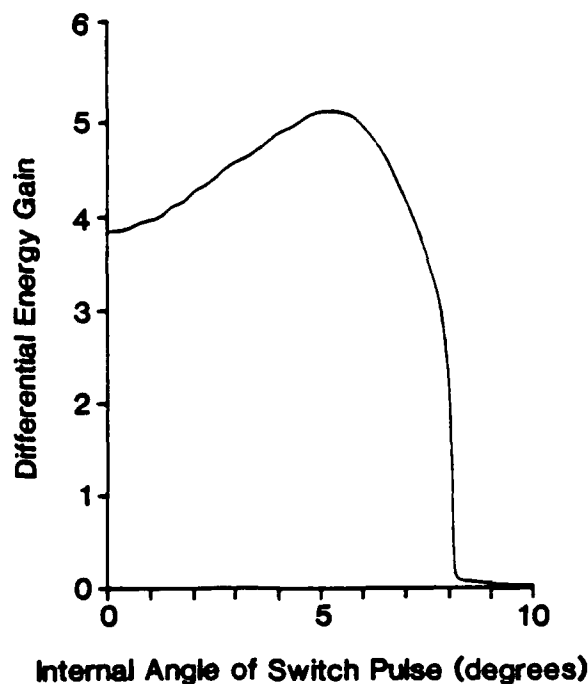


Figure 7.2 Gain vs θ for $\lambda = 883.6 \text{ nm}$, $I_{\text{pump}} = 17.1 \text{ kW/cm}^2$, $I_{\text{switch}} = 0.5 \text{ kW/cm}^2$, $\tau = 20 \text{ } \mu\text{s}$, $L = 2.042 \text{ } \mu\text{m}$.

7.4 Effect of Angle-Tuning on Switching Intensity

In addition to making the switching pulse more effective by allowing it to operate in a high-transmission frequency regime of the resonator, angle-tuning also allows a smaller switching pulse intensity to accomplish the needed shift of the Fabry-Perot peak. This reduces the E_s needed and provides an alternative source of increased gain, as shown in Fig. 7.4. Here curve 1 shows the behavior of the gain as a function of switching pulse intensity for normal incidence. In this case an intensity of approximately 400 W/cm^2 is required to switch the device optimally. Curve 2 displays the corresponding behavior for $\theta = 5^\circ$; a switching pulse intensity of only 200 W/cm^2 is required, and the gain is enhanced by nearly a factor of 2 over the case where $\theta = 0^\circ$.

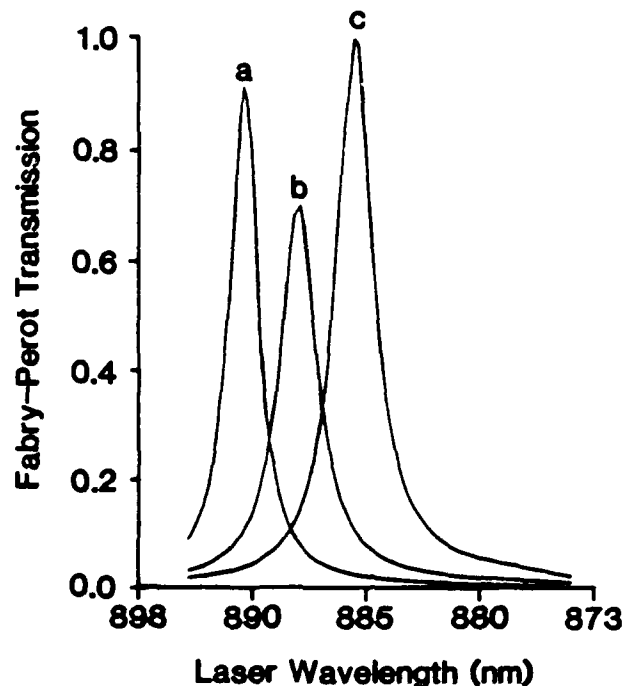


Figure 7.3 Shift in normalized Fabry-Perot peak as non-equilibrium charge carriers are excited. Curve a: $N = 10^{15} \text{ cm}^{-3}$; curve b: $N = 8.5 \times 10^{17} \text{ cm}^{-3}$; curve c: $N = 12 \times 10^{17} \text{ cm}^{-3}$. Increasing the carrier density shifts the peak towards the laser frequency, assumed to be at $\lambda = 883.6 \text{ nm}$. A switch pulse angle of $\theta = 5^\circ$ shifts the laser line effectively to 885.9 nm , where peak c is located.

7.5 Assessment of Plane-Wave Approximation and Beam Walk-off

The model for the dynamic behavior of semiconductor etalons producing the results discussed in this paper does not take into account the characteristics of finite-width beams. The potential effect of beam walk-off on the angle-dependent behavior of the etalon has been ignored also. We now discuss these two important topics, under experimental conditions for which the predictions made in the preceeding sections are likely to be achieved.

a) Plane-wave approximation

For the plane-wave approximation to be valid, a necessary condition is $Z_R \gg L$, where Z_R is the Rayleigh length of the Gaussian input beam, and L is the cavity length. Z_R is given by [7.10] $Z_R = w^2/\lambda$, where w is the beam waist radius and λ

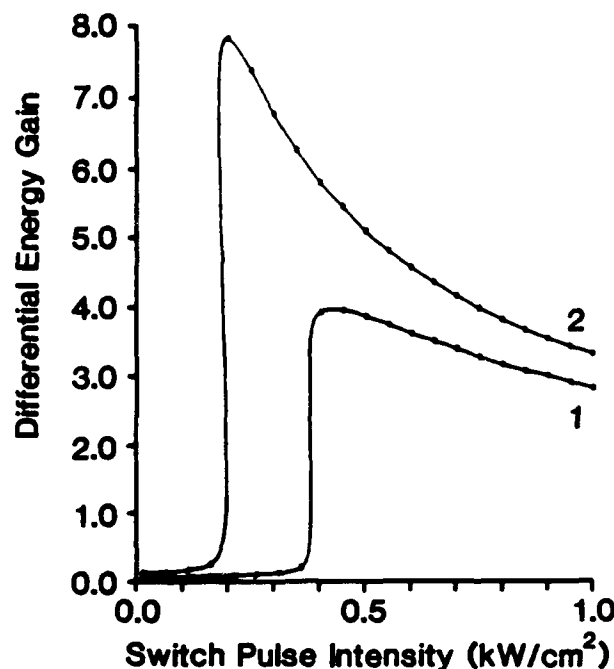


Figure 7.4 Comparison of gain vs switching intensity behavior for $\theta = 0^\circ$ (curve 1) and $\theta = 5^\circ$ (curve 2).

is the wavelength. If we consider typical experimental conditions for which $w = 5 \mu\text{m}$ and $\lambda = 0.884 \mu\text{m}$, then $Z_R = 25 \mu\text{m}^2 / 0.884 \mu\text{m} \cong 28 \mu\text{m}$. For $L \cong 2 \mu\text{m}$, then, $Z_R \gg L$ is satisfied.

A more rigorous formulation requires $Z_R > mL/\cos\theta$, where m is the effective number of times the light bounces within the cavity. Here m depends on R (the mirror reflectivity), α (the absorption coefficient), and L . θ is the angle between the switching and pump beams in the cavity, which, when non-zero, tends to increase the effective length over which absorption can take place. We will see in the next section that m is usually less than 10, implying $mL \cong 20 \mu\text{m}$ and allowing $Z_R > mL/\cos\theta$ to be satisfied for small θ . We conclude that the plane-wave approximation assumed in this paper holds under the experimental conditions anticipated.

b) Beam walk-off

Intuitively, the beam walk-off problem is minor if the beam spot size is much greater than the etalon thickness, and if the angle of the beam is not too large. The

amount of walk-off should also depend on the absorption inside the cavity and on the reflectivity of the mirrors. A good limiting condition is [7.8]:

$$\mathcal{F} = \frac{n_0 \pi w^2 T}{\lambda L} \gg 1 \quad (7.5)$$

where \mathcal{F} is the Fresnel number, n_0 is the refractive index, and T is the mirror transmission. Inserting the values $n_0 = 3.514$, $w = 5 \mu\text{m}$, $T = 0.1$, $\lambda = 0.884 \mu\text{m}$, and $L \cong 2 \mu\text{m}$, gives $\mathcal{F} \cong 16$, which means that the beam walk-off is probably minor.

A similar approach to the problem is to estimate m , the effective number of times the light bounces within the cavity. One can define m in the following manner:

$$E_0 R^m e^{-\alpha L m / \cos \theta} = E_0 e^{-2} \quad (7.6)$$

where E_0 is the initial amplitude of the light field. If we take the parameters used in the simulations, namely, $R = 0.9$, $\theta = 5^\circ$, and $\alpha L = 0.24$, then $m \cong 6$. Even if $\alpha L = 0.1$ is assumed, we find $m \cong 10$. The lateral shift of the beam after 10 bounces is about $3.5 \mu\text{m}$. Because the spot diameter is $10 \mu\text{m}$, most of the light remains within the interaction region of the two beams. So, we expect the beam walk-off problem to be minor for the conditions considered here.

It should be noted that the angle-tuning technique described in this paper may not be effective if the light is too tightly focused, as this will result in a small spot size (with respect to the cavity length) and a large divergence angle (i.e., a short Rayleigh length).

7.6 Conclusions

In dynamic simulations of bistable switching behavior where two beams address a GaAs etalon, non-normal incidence of either the signal or switching beam has been found to enhance the gain achievable in such devices by a factor of two for single-wavelength, μs pulses. Adjusting the angle between the two beams leads to an increase in the amount of the pump beam which can be transmitted effectively, or to a decrease in the signal intensity required for switching to occur. Although plane-wave behavior is assumed in all of these simulations, these results are expected to hold for finite-width beams as well. Further optimization of the cavity could result in greater enhancement of the gain. Although similar enhancement effects are expected for faster pulses, significant gain likely will remain infeasible for pulsewidths approaching the carrier lifetime of the material used.

8. FAN-IN OF SIGNALS TO A GaAs ETALON DEVICE

8.1 Introduction

To be useful, an optical logic or switching device must use two or more input signals to determine an output signal. A single output signal may be fanned-out efficiently into multiple signals. When N beams are fanned in to a single device, however, the efficiency is, in most cases, reduced to $1/N$. This loss in efficiency arises from interference between the beams to be fanned in if they are not precisely phase matched, whether they are coherent or not [8.1]. There are three exceptions to this occurrence: 1) when two beams of crossed polarization are combined; 2) when coherent beams to be combined have certain precise phase relationships [8.2]; and 3) when it is acceptable that the input power be distributed over an area equal to the total cross-sectional area of the beams to be fanned in. The third condition follows from the constant radiance theorem and also can be expressed in terms of the numerical aperture or acceptance angle of the system [8.1].

A way must be found to cope with these losses, or to use the above exceptions to avoid the losses. Exception 2) usually is not practical. Exception 3) may be useful when the diffusion distance in a pulse length or carrier lifetime (whichever is shorter) exceeds the spot diameter, since in that case the light beams may not be required to overlap. Exception 1), combining two beams of crossed polarization, is effective, but this method is limited to a fan-in of two beams with low loss. This is the method we will describe in the rest of this section. We attempted two other methods to reduce losses to interference: 1) we tried bringing two focused spots adjacent to each other, but not quite overlapping, so that they would interact with each other through diffusion of the carriers in the GaAs, without suffering interference effects; 2) we tried bringing two overlapping spots in at different angles of incidence, to look at switching as a function of the angles between the beams. With both methods we ran into experimental difficulties. We were unable to create a satisfactory optical system for producing two nearly-overlapping focused spots.

In order to get ~ 15 nm diameter focused spots, we used a $10\times$ microscope objective, and were producing two nearly-parallel beams, closely spaced, by reflecting one beam off a mirror. However, these beams could not be brought close enough together to fit simultaneously through the lens, nor could they be brought close enough to parallel to be nearly overlapping. We considered the following solutions to these problems: 1) Use a beamsplitter to combine the beams. We would have lost

one half of the input light, and did not have enough laser power available to follow this option. 2) Combine two orthogonally-polarized beams with a polarizing beamsplitter. This method works for a maximum of two beams. 3) Use a larger lens. The microscope objective already has a very low $F/\#$; a larger lens with the same focusing power probably would be difficult to obtain.

We had similar difficulties when we tried to combine two beams brought in at different angles.

Any system we made in which the spots overlapped had such large spot sizes that there was insufficient light intensity at the center to produce switching in the etalon. This might not have been a problem if we had been working with a more powerful laser.

Thus the main problems we encountered did not involve interference of light beams but with system geometry. Lenses large enough to admit two beams did not have a focal length short enough to produce a spot small enough to make the etalon switch. When we tried using a small lens for each beam, we had to choose lenses with focal lengths great enough to allow us to fit the lenses close to the sample and still see the beams exit. The resulting spot size was not small enough to produce switching with our laser. The problem of combining the input beams without a loss of intensity turns out not to be trivial.

8.2 Experiment

Our optical system for fan-in with crossed polarizations is shown in Fig. 8.1. Light from a single dye laser is divided into three beams using half-wave plates and polarizing beamsplitters. A strong beam is used as a holding beam to bias the GaAs etalon near its switching point. Two weaker beams are used as signal beams, representing the input of data into the device. In this design, the two signal beams are polarized in the same direction and converge on the etalon from opposite directions. The holding beam is polarized opposite to the signal beams, and is combined with one of the signal beams through use of a polarizing beamsplitter.

The advantage of this system is that it allows the output signal to be separated completely from the signal beams and their back reflections. The output signal consists of light from the holding beam which is either transmitted by the device or not, depending on its state. Thus it can be separated completely from the signal beams and their back reflections by isolation of light of the desired polarization, through use of a polarizer or polarizing beamsplitter.

We tested this system as an optical AND gate. The etalon we used contained a sample of GaAs/AlGaAs multiple quantum wells, with 63 periods of 299-Å GaAs.

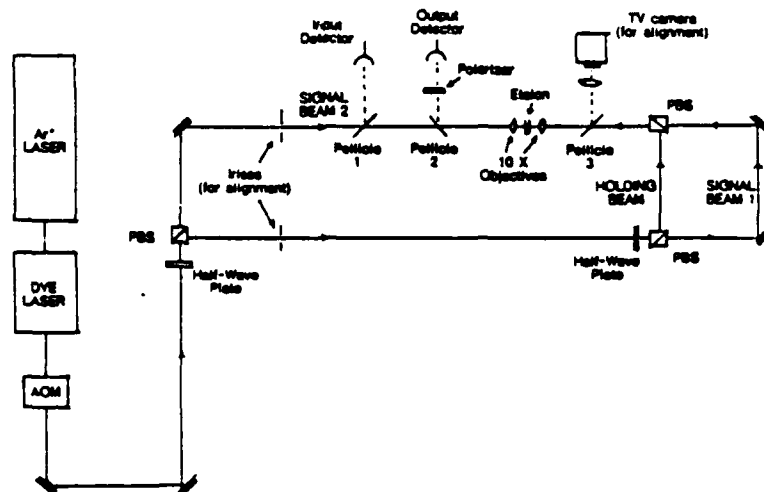


Figure 8.1 System for fan-in of three beams along a single axis, used to produce results seen in Fig. 8.2. Two signal beams of the same polarization approach the etalon from opposite directions. The holding beam is polarized orthogonally to the signal beams and is combined with signal beam 1 through use of a polarizing beamsplitter. If we replace pellicle 2 with a polarizing beamsplitter, the entire output signal can be isolated.

operating at a wavelength of 869.8 nm. The 299-Å MQW here is the same sample we characterized earlier. We used this MQW because the etalon was fairly uniform. We probably could have gotten a better result with the 58-Å-MQW samples we used for the cascading experiment, but the 58-Å-etalon was highly nonuniform. Results of this experiment are shown in Fig. 8.2. Peak power of signal beam 1 was 5.0 mW; peak power of signal beam 2 was 3.9 mW; and the power of the holding beam was 34 mW. Contrast between the high and low states is about 6 to 1. A further test was performed to maximize the output power with respect to the signal-beam power. The most favorable result showed an output of 8.5 mW for inputs of 7.0 and 3.5 mW for the two signal beams, and of 47 mW for the holding beam. Thus, there is enough power in the output beam to switch 1 or 2 identical gates. The asymmetry in the input power required suggests that further improvement may be made. An earlier result exhibited symmetry: inputs of 7 mW for each input

beam, with 39 mW in the holding beam, resulted in an output of 8.5 mW. In the former case, the stronger (47 mW) holding beam evidently allowed the use of lower signal-beam power. Possible improvements include the use of narrower quantum wells (50- to 100-Å layers instead of 299-Å GaAs layers) and a more uniform sample.

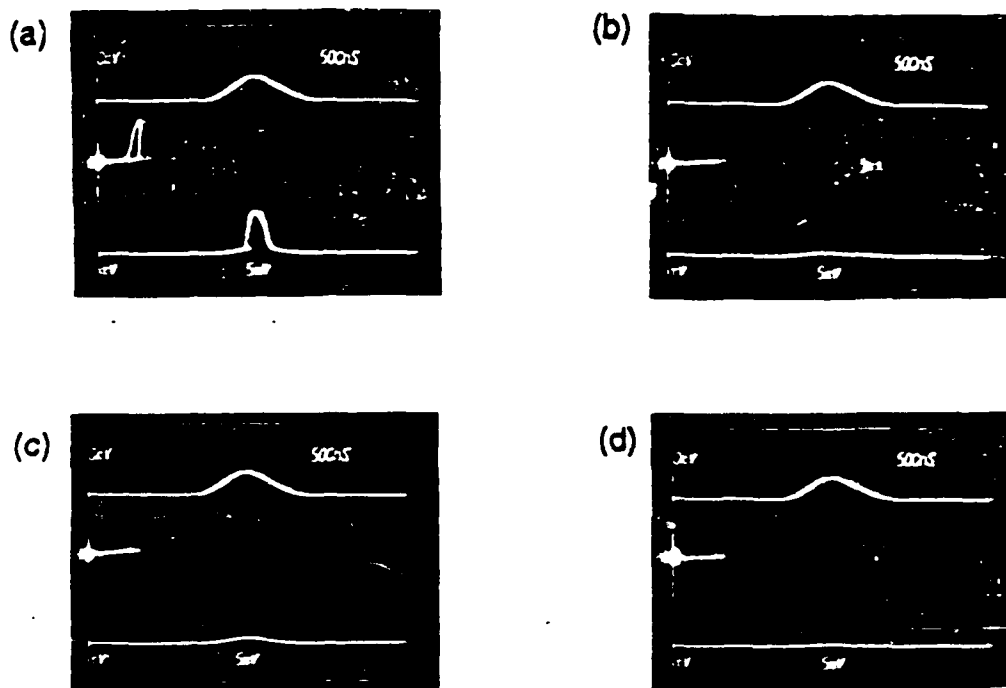


Figure 8.2 Demonstration of AND gate. (a) Output signal with all inputs ON. (b) Background from signal beams with holding beam blocked. (c) Output signal with signal beam 1 OFF, signal beam 2 ON, holding beam ON. (d) Output signal with signal beam 1 ON, signal beam 2 OFF, holding beam ON. In each figure, the upper trace is the input signal and the lower trace is the output signal, with output versus input displayed in the center. Signal beam 1 has a power of 5 mW, signal beam 2 of 4 mW, and the holding beam 34 mW. The wavelength of each beam is 870 nm.

Thus we have demonstrated efficient fan-in of two signal beams and a holding beam on a GaAs etalon optical AND gate operating at a single wavelength. Even though all beams are along a single optical axis, the output signal may be separated completely from the input signals because the holding beam is polarized orthogonally to the signal beams. The output power is great enough that the system should be cascable, which is consistent with the results of Section 7.

9. DIRECT FIBER-ETALON-FIBER INTERFACING

9.1 Introduction

The optical systems for cascading and fan-in shown in Figures 5.4 and 8.1 require much space and are sensitive to misalignment. One approach we have considered to make optical systems with GaAs etalons more compact and rugged is to connect them with optical fibers.

The main application of optical fibers has been as carriers of information in optical communication systems, because of their extremely high bandwidth. Full utilization of this high information-carrying potential awaits the development of high-speed opto-electronic or all-optical switching systems, so that the information, in the form of short light pulses, can be encoded, routed, and processed at a speed comparable to the fibers' bandwidth. Recently, the use of optical fibers as interconnects has attracted attention. An example is the use of an optical fiber to couple light into a waveguide [9.1].

The study of optical fibers used as interconnects in a digital optical system is increasingly important. For example, one might take advantage of the fast switch-on times of GaAs nonlinear etalons to achieve multi-GHz multiplexing and demultiplexing [9.2]. Previous experiments showed that all-optical data switching can be performed with GaAs etalons, with picosecond data pulses transmitted over a 1-km single-mode fiber, without affecting the device performance [9.3]. Here we demonstrate a technique that interfaces single-mode optical fibers directly with GaAs nonlinear etalons. In Section 9.2, we discuss the basic requirements of optical fiber interconnects, and present some preliminary experimental results with a GaAs/AlGaAs multiple-quantum-well (MQW) nonlinear etalon having reflective silver coatings. The improved uniformity of GaAs etalons with mirrors grown by MBE is discussed in Section 9.3. This improvement allows immediate extension of this technique to the fabrication of small arrays of optical fibers and optical logic gates. A more detailed discussion of the implications of this technique is given in Section 9.4.

9.2 Fiber-Etalon-Fiber Interfacing In Transmission

It is well known that in order to switch a nonlinear optical logic gate such as a nonlinear etalon, a *focused* spot of light is required to induce sufficient nonlinearity. If several such logic gates are to be interconnected, the output light from one logic

gate must be focused onto another. Usually lenses, prisms, mirrors and holograms are used as interconnection elements, but one often encounters problems in the implementation of a real system. These problems include restrictions in the volume of the interconnection elements; achieving appropriate spot sizes and uniformity; flatness of the focal plane; constraints in aberrations and beam divergences; the alignment, fan-in and fan-out of multiple beams; interference and cross-talk among different channels; timing of the light pulses; and the efficiency, cost, and ruggedness of a system. Considering these problems, conventional interconnection elements may not always be the best. Optical fiber interconnection, in combination with free-space interconnection, may prove to be advantageous in many optical signal processing and photonic switching applications, especially in those requiring a high data rate but not large scale parallelism [9.2-9.4] (Section 15). It is premature at this stage to give a more complete analysis of the potential of fiber interconnects. Initial results have shown great promise, however.

As stated earlier, the relative merits of fiber interconnects and of free-space interconnects are system-dependent; in most cases the optimum choice may be a combination of both. There are some basic system-independent requirements that must be satisfied for optical-fiber interconnects to be useful and competitive. We begin our research by studying some of these basic requirements:

- a) that a fiber can be interfaced directly to an optical logic gate, that is, that the output of a fiber can be used to switch the gate directly, without using other optical components to re-focus the beam;
- b) that a fiber collect the output of a logic gate easily and efficiently, so that the signal can be sent on to other logic gates without much loss;
- c) that the overall efficiency of the system be competitive. This requirement demands a high coupling efficiency between a light source and a fiber; comparable or lower switching power when the output is from a fiber, relative to that when using a focusing lens; and a high efficiency in coupling the output light from a logic gate into a fiber.

Our experiments are designed to test whether these basic requirements are satisfied by optical fiber interconnects. The optical fibers used in our experiment are from AMP Incorporated. The ends of these fibers are polished and mounted in ceramic housings. The fibers are 3 to 4 meters in length, and their core diameter is 10 μm . The light source is a CR590 dye laser pumped by an argon ion laser. The dye laser has a tunability from 815 nm to 905 nm. The laser beam is modulated by an acousto-optic modulator to give pulses of several microseconds in duration. The

repetition rate of the pulses is kept low to avoid etalon-heating effects. The laser light is coupled into the input fiber through use of a 10 \times microscope objective with a numeric aperture of 0.3. The output ends of the fibers are imaged onto an infrared camera to allow the corresponding modes to be examined with a TV monitor. The fibers are designed for optimum use in the 1.3- μ m wavelength region. To use the laser light efficiently, the relative lens-fiber position is such that the output from the fiber is as uniform as possible. The total loss (occurring mainly during coupling of the light into the fiber) can be limited to less than 3 dB with little effort. A nonlinear etalon is placed against the output end of the input fiber; a similar fiber is placed behind the etalon to collect its output. The experimental setup is shown in Fig. 9.1(a).

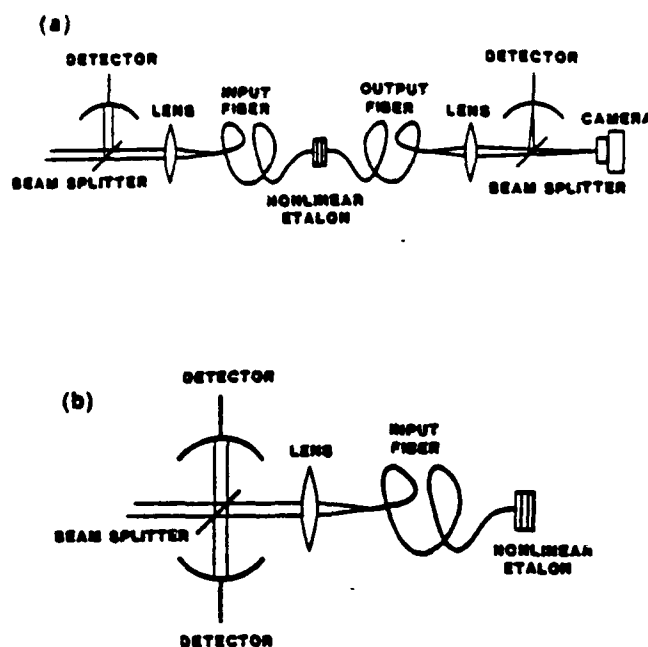


Figure 9.1 Experimental setup for interfacing fibers with a nonlinear etalon operating (a) in transmission mode and (b) in reflection mode.

Etalons made previously by sandwiching the nonlinear crystal between two dielectric mirrors (Section 4 and [9.5]) are not suitable for this experiment; the mirrors were deposited on glass cover-slips of a few hundred micrometers in thickness, making it difficult to collect the transmitted light. The nonlinear etalon

used in this experiment was made by coating both sides of a 180-period 58-Å GaAs/96-Å AlGaAs MQW crystal with an area of 2 mm², with silver 400 Å thick, resulting in reflectivities of about 85%. The 58Å-MQW sample was chosen because the etalon peak in this sample was near the wavelength of the peak power of our laser. Optical bistability is observed (see Fig. 9.2) at a wavelength of 842 nm when the etalon is sandwiched between the two fibers (Fig. 9.3). Although the beam diverges upon exiting the input fiber, its diameter is not much larger at the entrance of the output fiber. This is because the total thickness of the etalon is small (a few micrometers) and the refractive index of the etalon material is large. In fact, more than 70% of the light transmitted through the etalon is recaptured by the output fiber.

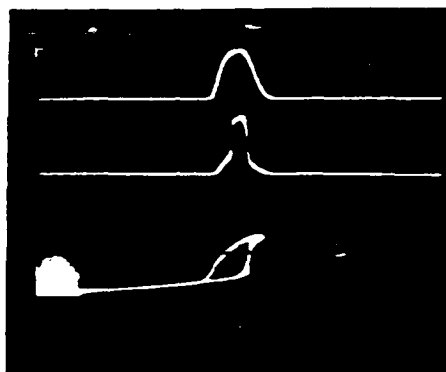


Figure 9.2 Optical bistability observed from the output fiber, with the etalon switched by the light from the input fiber directly (without using a lens to focus the beam). The input power is about 60 mW.

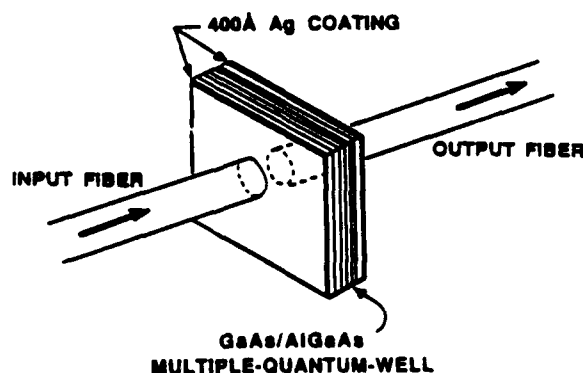


Figure 9.3 Interfacing technique used to bring light into and out of a nonlinear etalon with optical fibers as interconnects.

An interesting trade-off in switching-power results when using a lens-fiber combination rather than a lens alone to switch the etalon. In the experiment with the silver-coated etalon, the power required for observing optical bistability is about 100 mW if a 10 \times microscope objective is used to focus the light onto the etalon, *and results in a spot size of around 13 μ m*. If the same objective is used to couple light into a single-mode optical fiber (with a loss of less than 3 dB), and the output from the fiber is used to switch the same etalon, then the required switching power is only about 60 mW, but results in the same light intensity on the etalon. This reduction in required power, confirmed in experiments employing samples with integrated mirrors operating in reflection mode, will be discussed in more detail in the next section.

9.3 Fiber-Etalon Coupling in Reflection

Using Devices With MBE-Grown Mirrors

Etalons with highly uniform properties are desirable if one is to operate a large number of optical logic gates under similar conditions, as in an array to be used for parallel processing. The nonuniformity of the etalons made previously [9.5, 9.6] is caused mainly by variations in the thickness of the optical adhesive used to glue the dielectric mirrors onto the nonlinear crystal, and to the wedges that are introduced during the chemical-etching process. The uniformity of an etalon may be improved by growing mirrors on the sample when it is fabricated, rather than adding the mirrors to the etalon after fabrication of the nonlinear optical material.

Such an etalon was designed and grown by MBE (see Section 4). For simplicity, in our first attempt at fabricating an etalon with integrated mirrors we specified a nonlinear layer of bulk GaAs 1.6- μ m thick. The mirrors are fabricated from alternating layers of AlAs and GaAs *forming quarter-wave stacks* on a substrate of semi-insulating GaAs, as shown in Fig. 4.8. The uniformity is greatly improved over the etalons made previously by adding mirrors after growing the nonlinear crystal [9.5, 9.6]. The finesse of the etalon is about 20, with a 2.1-nm linewidth (FWHM). The peak wavelength of the Fabry-Perot cavity is located at 891.1 nm in the center of the wafer. Over a 1 cm² area at the center of the wafer, the Fabry-Perot resonance shifts monotonically by only three linewidths (see Sections 4 and 12). The shift is even less in a smaller region; for a half-circle of radius of about 6-mm around the center, the peak is shifted only one linewidth. Optical bistability can be seen at a wavelength of around 888 nm by focusing the beam directly on the etalon.

The experimental setup for fiber-etalon-fiber interfacing is shown in Fig. 9.1(b). The light reflected from the etalon is collected by the same fiber that transmits the

incident light, and a beamsplitter is used to divert a portion of the reflected light into a detector. Figure 9.4 shows the bistability curve of the light reflected back through the fiber. In this experiment more than 50% of the reflected light came back through the fiber and the lens, and the beamsplitter reflected about 7% of that light into the detector. By using a polarization-preserving fiber or an ordinary fiber with a short length (so that most of the polarization is maintained) in combination with a polarizing beam splitter and a quarter-wave plate [9.5] (Section 5), it is possible to collect most of the reflected light. Operating the etalon in reflection mode alleviates the problem of fiber alignment in fiber-etalon-fiber coupling, as the output fiber is also the input fiber. If thermal effects degrade the operation of an etalon based on an electronic nonlinearity, one side of the etalon may be mounted to a heat sink when the etalon is operated in the reflective mode.

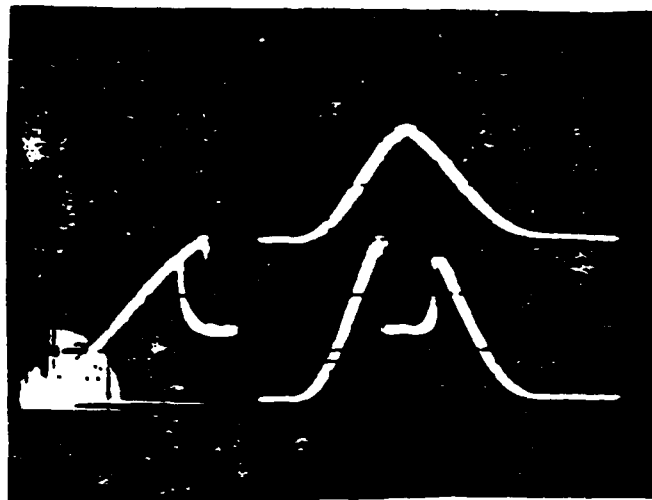


Figure 9.4 Optical bistability seen in the light reflected back into the fiber, corresponding to the setup shown in Fig. 9.1(b). The switching power is about 26 mW. The contrast between high and low states is about 3:1 when background reflections from the optical system are subtracted.

The switching-power trade-off seen in transmission mode is seen also in reflection mode. The etalon was switched with 26 mW when it was interfaced with the fiber, whereas 42 mW was required when light was focused directly on the etalon using a 10 \times microscope objective, giving a spot size of about 13 μ m. Measurements for use of a 16 \times microscope objective suggest that there is probably a reduction in switching power with a smaller spot size. However, use of a smaller apparent spot size will not always reduce the switching power, because the effective spot size in the time regimes of the experiment can be increased significantly by

diffusion of the photo-generated carriers. It has also been shown that using the *lower-order* modes of the fiber requires less switching power than does using the higher order modes, since the corresponding intensity distribution is more favorable. This means that a fiber could act like a spatial filter that cleans up the initially aberrated wavefront, thus giving a lower switching power. A more detailed study is required to determine the combined effects of the mode quality and the optimum spot size (which could be diffusion-limited) on the switching power. When a 10 \times objective is used to couple light into the fiber the coupling loss is less than 3 dB. Further improvements on the coupling efficiency are expected with a better match of the numerical aperture of the focusing lens and the modes inside the fiber.

9.4 Discussion

The experimental results described in Sections 9.2 and 9.3 show that fiber interconnects satisfy the basic requirements listed in Section 9.2. The fibers can be interfaced directly to optical logic gates (nonlinear etalons in our case), and the output from a gate can be collected efficiently by a fiber (a 50-70% efficiency has been demonstrated). The overall system efficiency is also competitive; light can be coupled into a fiber with a loss of less than 3 dB without much difficulty, and the switching power using the output light from a fiber can be lower than that for light focused by a lens that couples it into the fiber. Further improvement in system efficiency is achievable through use of single-mode fibers designed to operate in the 850-nm wavelength region. An example is the experimentally demonstrated coupling efficiency of 75%, using a 40 \times objective lens to couple light into a 5- μ m-core fiber [9.3]. A reduction of switching power is also expected with a smaller core size using our technique (which may be difficult if a lens is used to re-collect the output light from the fiber, resulting in losses and aberrations), and the fibers can easily be glued to the etalon to improve the stability. The uniformity of the GaAs etalon with mirrors grown by MBE allows immediate extension of our current results to fiber arrays. This technique also is easily extendable to similar devices, such as InGaAs/InP-MQW etalons that operate in the 1.55- μ m wavelength region [9.7].

The potential for optical fiber interconnections in an optical processor depends—critically—on the architecture of the processor. In general, optical processor architecture can be divided into two categories: serial and parallel. An example of high-bandwidth serial processing is the photonic switching used in optical communications, where a large amount of information is multiplexed into a few channels. Suppose an optical processor (such as a demultiplexer) made with optical switches (such as nonlinear etalons or waveguides) is used to process the information.

Because the light signal is already inside a fiber in this case, the simplest interconnection scheme may be to bring the light from the fiber to the optical logic gates directly, using fiber interconnects. For example, in the all-optical compare-and-exchange switches discussed in Section 15, some lenses and prisms can be replaced by fibers. Pulses from different sources can be combined by using fiber couplers, and their timings can be adjusted easily by choosing the appropriate fiber length (an accuracy within a few picoseconds is relatively easy to achieve). In these processors, the use of fibers may make the whole system more efficient and more rugged, and may also reduce the volume and cost. Optical fibers may even be used as an efficient interface between different types of optical switches, such as nonlinear etalons and nonlinear waveguides.

Optical fiber interconnection may be able to preserve its advantage even with a limited amount of parallelism. For example, significant enhancement of system capability could be achieved with a moderate extension to a 10×10 fiber array, which would give 100 parallel channels. What fiber interconnection can offer, in addition to the advantages discussed above, is: the fixed spot size; the uniformity of the spot size and the light level for each spot; the extremely low cross-talk between the channels; ease of control of beam divergences (especially for beams coming from different sources); and the fixed angle of incidence. However, the extension of fiber interconnects to applications that require large-scale parallelism (such as the parallel operation of 10^4 to 10^6 pixels on an etalon) does not seem feasible at this stage, because coupling light into a large number of optical fibers is difficult. Another possible disadvantage for fiber interconnection is its limited packing-density; it is difficult to obtain inter-pixel separations of less than $100 \mu\text{m}$ because of the cladding size. Free-space interconnection is clearly preferred in this case. A combination of free-space and fiber interconnections may be the best choice for some applications.

9.5 Conclusions

We have demonstrated that a nonlinear etalon can be interfaced directly and efficiently with optical fibers. This result may lead to a number of interesting developments, such as the 2-D and 3-D integration of fiber interconnects and optical logic elements. The use of fibers with multiple inputs and the availability of various kinds of fiber couplers will allow fan-in and fan-out of more beams. One important advantage of optical fiber interconnects is that they are system-dependent; they may play an important role in a high-speed or complex digital-optical system.

10. HOLOGRAPHIC OPTICAL INTERCONNECTS [10.1]

10.1 Introduction to Optical Interconnects

This section addresses issues specifically related to optical interconnection. Two optical-interconnect applications will be discussed. The first is the case for an optical computing system in which optical interconnects are an essential element for implementing optical-computing architectures. For example, a globally connected neural network system requires that each neuron be connected to other neurons with dynamical interconnections [10.2, 10.3]. Other applications of regular optical interconnects can be found in digital optical computing, as described in Section 13 [10.4, 10.5].

The second application of optical interconnects is the connection of electronic integrated circuits (ICs). Photonic interconnect techniques are necessary in part because of the trend to reduce the size of electronic devices. Other related problems are the power requirements for output lines, interconnect complexity, time delay, and transmission line mismatch.

The development of ultrashort pulse generation, e.g. as short as $\cong 8$ fs by optical techniques [10.6] and ultrafast switch-on times of all-optical bistable logic gates, e.g. several picoseconds in GaAs etalons, indicated a computation potential that is several orders of magnitude faster than electronic counterparts. It also permits the evaluation of systems limited by photon rather than electron transport mechanism.

To achieve both fast processing and large optical communication bandwidth simultaneously, intermediate optical-to-electrical and electrical-to-optical conversions must be eliminated. This can be accomplished by using optical gates in conjunction with optical connections of these gates. Therefore, a driving force for optical interconnection here is to provide an environment as optical as possible, so that all computation can be performed optically to fully explore the limits of the optical environment. As described in the demonstrations of optical computing (Section 13), utilization of classical optics and diffraction optics is clearly appropriate for implementing free-space optical interconnects for an array of all-optical logic gates.

The optical-interconnects techniques employed for optical computing can also be extended to solve interconnection problems in VLSI technology. Both applications can take advantage of the intrinsic properties of optics: speed, parallelism, and interference immunity. For example, both the interconnection among neurons and the clock signal distribution among chips require some or all of these properties. In the

remainder of this section, some problems of electrical interconnects are discussed, and the use of optical interconnection as an alternative solution to these problems is addressed.

Figure 10.1 shows the history of the scale of integration for silicon ICs [10.7]. The progress made in the very large scale integration (VLSI) technology suggests that electronic-system performance can be improved and the cost reduced by developing smaller devices, and by using only a few ICs with faster device speeds and lower power requirements, to perform the functions of many ICs. However, this trend cannot continue indefinitely. Two factors that limit the reduction in scale are the device density per unit area and signal propagation delays between devices and integrated circuits.

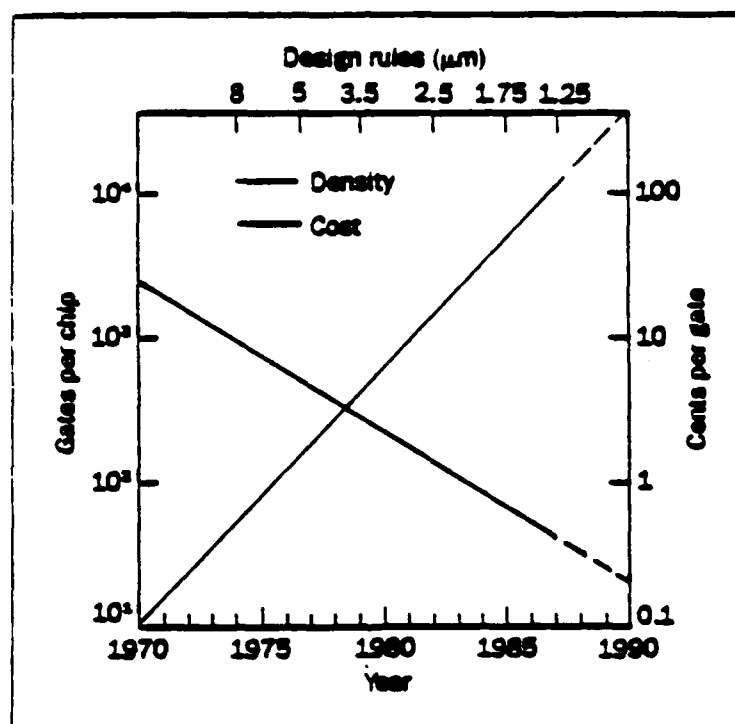


Figure 10.1 Trend of VLSI capability (from AT&T Technical Journal, 1987).

(I) Physical Space Limitation

It is noted from Fig. 10.1 that the maximum number of components per chip has increased approximately by a factor of 100 in the last ten years. As the number of devices increases, the number of output/input connectors and intrachip interconnections also must increase. One relationship that describes the corresponding

number of input/output pins and the number of gates on a chip can be estimated by Rent's rule [10.8, 10.9],

$$p = kg^a, \quad (10.1)$$

where p is the number of pins per package, g is the number of gates in the package, k is a parameter in the range from 3 to 6, and a is from 0.45 to 0.55 for random logic systems. A typical chip with a 1-cm^2 area can accommodate 8000 gates, requiring 170 to 840 pins. Each pin needs a bonding pad that occupies an approximately $100\text{-}\mu\text{m}\text{-x-}100\text{-}\mu\text{m}$ area.

It is also noted from Fig. 10.1 that the size of a conventional logic device has decreased approximately by a factor of 10 in the past twenty years. A factor of about three to five smaller will produce a minimum channel length of ≈ 0.25 to $0.5\text{ }\mu\text{m}$ for a silicon MOS device [10.10, 10.11]. Further reduction in size is not expected to produce useful devices.

As circuits are scaled down, more and more devices are packaged together over the same area. Therefore the spacing between lines decreases, and the density of electrical interconnects increases. The scaling effect brings about several problems, such as RC time delay, crosstalk, and heat dissipation. RC time delay is associated with electrical interconnects and can be a significant portion of the total circuit delay for VLSI [10.12]. Crosstalk is attributable to the capacitive and inductive coupling between two parallel electric lines whenever they are placed close to each other. As the line spacing is reduced, the capacitive crosstalk, which is also a sinusoidal function of line length, will dominate [10.13]. The heat-dissipation problem is the result of power losses encountered when switching large numbers of devices per unit volume. This is further aggravated by an increase in the speed of logic devices and will be discussed later.

(II) Propagation Time Delay Limitation

When an electrical signal propagates along a wire, the propagation speed is equal to $c/\sqrt{\epsilon_r}$, where c is the speed of light and ϵ_r the dielectric constant of the insulator surrounding the wire. For a fiberglass PC board, the dielectric constant is approximately 4; therefore signals travel at the speed of about 15 cm per nanosecond. At high computational speeds, the rise time of a signal can be subnanosecond; therefore the interconnection length needs to be at most a few tens of centimeters to prevent propagation delay. However, in general, on-board interconnect line lengths can vary from less than a centimeter to almost one meter, which imposes limits on

computational speed. Therefore, when designing microelectronic packaging, it is critical to minimize wire lengths and to maximize circuit density. As a rule of thumb for the design, the length of interconnection should be comparable to the physical distance that a signal itself occupies so that the interconnection delay becomes negligible. The condition can be expressed as

$$L \text{ (cm)} \leq v\tau_r = \frac{30(\text{cm/nsec}) \tau_r(\text{nsec})}{\sqrt{\epsilon_r}} = \frac{30p}{\sqrt{\epsilon_r} B} \quad (10.2)$$

where v is the signal propagation speed along a wire; τ_r is the pulse rise time, which occupies a small fraction p of the pulse and p is generally required to be less than 0.15 for logic operation; the bit rate B is the reciprocal of pulse duration. For example, if the maximum clocking frequency of a silicon circuit is about 1 GHz, interconnection lengths should be less than 90 mm. In a GaAs circuit operating at 10 GHz, interconnect length should be less than 16 mm. Therefore, electrical connection lengths of a few millimeters can pose a potential delay problem.

Referring to Eq. (10.2), one may consider the material of a dielectric constant lower than silicon, e.g. polyimide ($\epsilon_r \cong 3.2\text{--}3.8$) as an alternative for signal transmission media because, for a given τ_r , lower values of ϵ_r allow longer interconnection lengths. The improvement by reducing the dielectric constant is not significant because the dielectric constant is frequency and temperature dependent. In the extreme case, the maximum improvement from this idea will be a factor of two when $L \leq c\tau_r$, that is, $\epsilon_r = 1$. But this is not possible because the signal is then travelling in vacuum and not in a wire.

Another alternative might be to utilize a superconductive material, which offers the following advantages for interconnects: negligible loss, higher phase velocity, virtually nonexistent dispersion. However, the integrated capability of superconductive materials with silicon or gallium arsenide circuits and the termination procedures are unknown. In addition, room-temperature operation has not been achieved yet.

Problems also occur when high-frequency signals propagate along interconnects of a highly dense circuit. An electrical wire can be treated as a simple conductor; however each interconnect in a highly dense circuit cannot. It should be treated as a distributed transmission line [10.14]. In addition, the layout of electrical interconnections often requires that lines be bent and/or split, causing impedance mismatches at junctions. The reflection coefficient of a capacitive mismatch C can

be approximated by $\rho_C \cong R_L C / \tau_r$, and by $\rho_L \cong L(R_L \tau_r)$ for an inductive mismatch, where R_L , C , and L are resistance, capacitance, and inductance of a transmission line respectively, and τ_r is the rise time of a pulse [10.15]. Clearly, the power that results from impedance mismatch will increase with high frequencies. In packaging, it is therefore essential to have well designed interconnects, matching terminations and connectors to reduce to an acceptable level the effects of reflected signals, crosstalk, and signal attenuation.

Heat removal is one challenging problem in the packaging design. Interconnection lines of electronic systems must be charged and discharged, giving a dynamic switching power, which can be expressed as

$$P_{\text{line}} \cong (1/2) C_{\text{line}} L V_{\text{logic}}^2 f, \quad (10.3)$$

where C_{line} is a line capacitance per unit length, L is line length, V_{logic} is the logic voltage swing, and f is the clocking frequency. When the clocking speed is high, from Eq. (10.3), the heat generated per unit volume becomes large, and the thermal problem becomes even worse as the packaging density is increased [10.8]. In addition, line termination is necessary when line length is longer than the condition of Eq. (10.2), which requires relatively high-power consumption [10.16].

Optical interconnections, extensively used for long-distance telephone communications, have also been adapted for use with short-distance local area networks, and for machine-to-machine computer connections. At the machine-to-machine level, the use of optical fibers rather than the electrical cables has become a trend because of several overwhelming advantages: low crosstalk (thus less coaxial sheath); low attenuation per unit line length (thus fewer repeaters); high bandwidth and low dispersion (preserving sharp signal transitions); and immunity to electromagnetic interference (EMI). Consequently, optical fibers are widely applied to provide superior performance and cost-effective interconnections.

At the chip-to-chip level and the board-to-board level, it was reported that optical interconnects have the potential to overcome the previously mentioned limitations of electrical interconnects [10.17, 10.18]. Difficulties with optoelectronic packaging remain an issue, and it is difficult to determine the lowest level at which optics will make an impact. It is expected, however, that several advantages can be offered by replacing certain critical electrical transmission lines with optical interconnects: 1) fewer layers of metallization and polysilicon on the chip; 2) faster and larger fan-out communication with much less clock skew; 3) fewer electrical pins per chip; 4) direct access to locations away from the edges of the chip for testing; 5) capability of reprogrammable interconnections.

Waveguide, fiber, and free-space interconnections have been proposed at the chip-to-chip level. Interconnects using only optical fibers are not practical, because the connections consume too much space on an IC chip. Research on waveguide couplers using integrated prisms or gratings will help to couple modulated signals from laser diodes into on-chip waveguides for near-term chip-to-chip optical interconnections. However, in all cases a waveguide must be integrated and constrained to the IC plane. Waveguides also suffer from bending loss and coupling loss.

An alternative to guided-wave optics is free-space interconnects. This alternative offers potential advantages over guided-wave interconnects: 1) freedom from the planar constraint; 2) potential increased device density because space previously dedicated to waveguide or electrical interconnects on a chip is available, and 3) considerable freedom in designing the interconnection pattern and in reprogramming the pattern.

To achieve the above advantages, holographic optical elements (HOEs) should fulfill the following requirements: i) high efficiency, to reduce source power consumption; ii) capability of generating tightly focused spots at the detector area; iii) capability of generating arbitrary interconnect patterns; iv) reasonable misalignment tolerance to facilitate automated assembly; v) reasonable wavelength tolerance to accommodate the spectral distributions that are characteristics of laser diodes.

Computer-generated holograms (CGHs) are considered very general elements for producing arbitrary wavefronts of even nonexistent objects. This capability is very appropriate to provide the flexibility required for interconnects. However, CGHs generated by common plotters often suffer from two limitations, namely low-diffraction efficiency and little separation of the desired diffraction output from strong dc background caused by the small space bandwidth product of the plotters. The use of electron-beam lithography to write a hologram can significantly increase space-bandwidth product [10.19]. Electron-beam techniques are, however, quite expensive and not easily accessible. Alternatively, optically recorded holograms can approach diffraction efficiencies near 100% in recording material such as dichromated gelatin (DCG). Therefore, the idea of using a CGH to generate desirable wavefronts and then recording them in DCG material is expected to combine the flexibility of CGHs with the high diffraction efficiency and reconstruction fidelity of DCG [10.20, 10.21].

In the remaining subsections the design technique for constructing a CGH is described. A binary phase grating, which is capable of generating multiple beams of equal intensities and controlled angular separations, is used as an example. These

phase gratings are potentially important for regular optical interconnects as will be seen from the experimental results presented later.

A review of the coupled wave theory for thick holograms is then given with emphasis on the experimental results. Finally, an experiment that copies a CGH onto a DCG is demonstrated; results and applications are discussed.

10.2 Binary Phase Grating as CGH

Surface relief gratings have been extensively studied by many researchers. Detailed mathematical treatments of surface relief gratings can be found in the literature [10.22, 10.23, 10.24]. For optical-interconnect applications, it is often desirable to have a device that can split a single input beam into multiple equal output beams. High diffraction efficiency is also required for efficient power distribution. A binary phase grating is capable of generating multiple beams with high diffraction efficiency [10.25, 10.26, 10.27, 10.28]. The diffraction efficiency (η) of a grating is defined as the total power diffracted into the desired image field divided by the total power incident on the grating. Diffraction efficiencies of over 80% or more can be achieved with some binary phase gratings [10.28]. A motivating factor for developing binary phase gratings is to use them as efficient multiple beam splitters to meet the need of optical interconnection in optical computing and to show the feasibility of implementing these elements as optical interconnects for VLSI technology. Other applications for binary phase gratings include multiple imaging, star couplers [10.26, 10.29], coherent summation of laser beams [10.30], laser beam profile shaping [10.31], optical testing, and innovative optical applications such as aspheric lens testing in the IR regime.

Design Work

Consider a binary phase grating with a periodically rectangular phase profile as shown in Fig. 10.2(a), where the phase angle has two discrete values $\pm\phi_0$. It is well known that a thin grating can diffract an incident beam into many orders [10.32]. For a normally incident beam, the diffracted beams are governed by the following grating equation

$$d \sin \theta_m = m\lambda, \quad (10.4a)$$

where d is the grating period; m is the diffraction order and θ_m is the diffraction angle of m -th order.

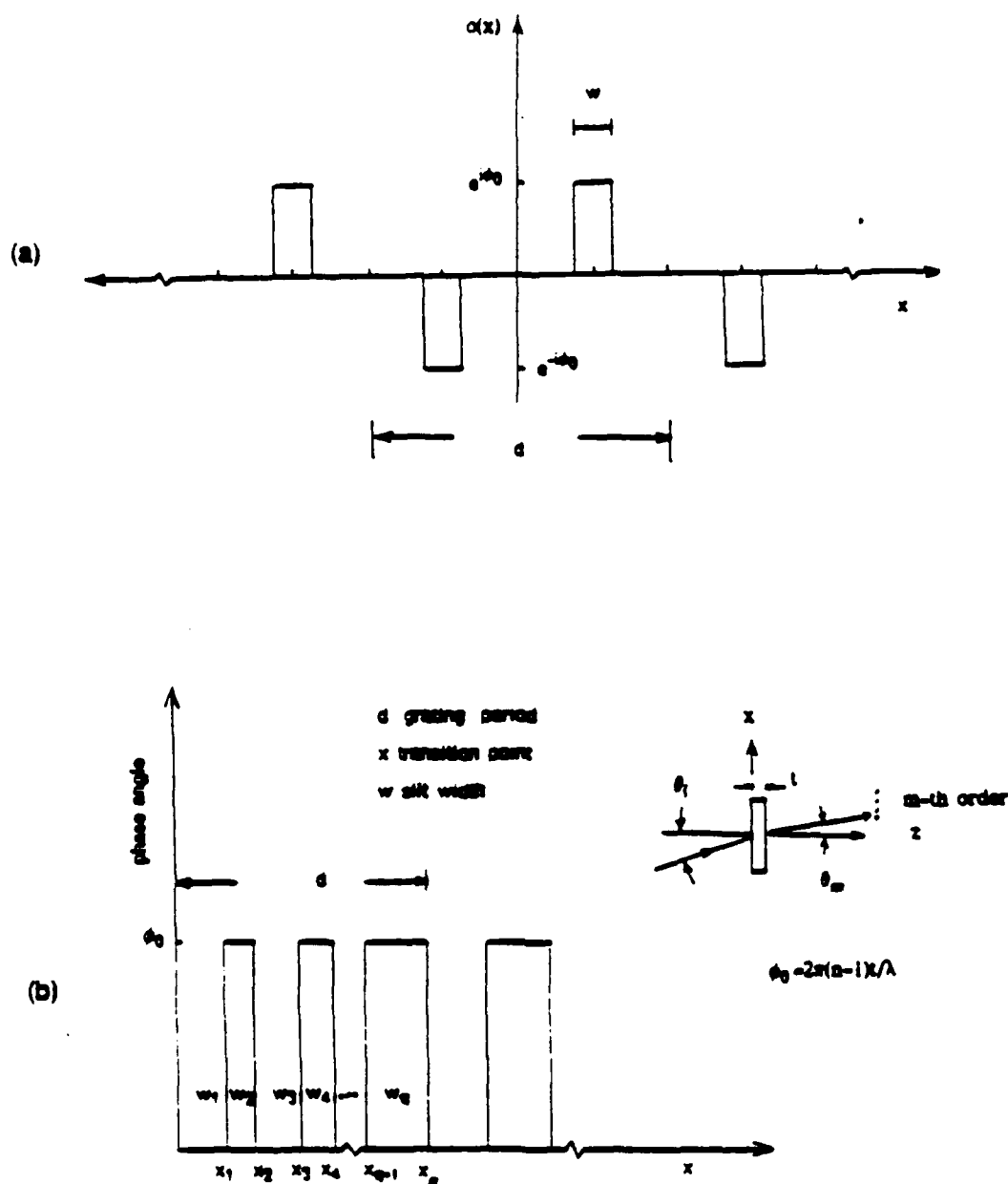


Figure 10.2 Binary phase grating with: (a) simple structure; (b) involved structure.

The optical transparency function of the grating $\alpha(x)$ shown in Fig. 10.2(a) can be written as

$$\alpha(x) = \left[e^{-i\phi_0} \text{rect} \left[\frac{x+d/4}{w} \right] + e^{i\phi_0} \text{rect} \left[\frac{x-d/4}{w} \right] \right] * \frac{1}{d} \text{comb} \left[\frac{x}{d} \right] \quad (10.4b)$$

where $*$ is a convolution operation and w is the slit width.

Assuming that the grating is illuminated by a plane wave, the amplitude of the m -th diffraction order in the far field can be calculated by taking the Fourier transform [10.33] of the transparency function $o(x)$ of the grating. Using Eq. (10.4a), the normalized Fourier transform amplitude in the far field is

$$\begin{aligned} O(\xi) &= \mathcal{F}\{o(x)\} = \int_{-\infty}^{\infty} dx o(x) e^{-i2\pi\xi x} \\ &= 2w \cdot \text{sinc}(w\xi) \cos(\phi_0 - \pi d\xi/2) \text{comb}(d\xi) \\ &= \frac{2\sin(m\pi(w/d))}{m\pi} \cdot \cos(\phi_0 + \pi m/2) \sum_{m=-\infty}^{\infty} \delta(\sin\theta_m - m\lambda/d), \end{aligned} \quad (10.4c)$$

where $\xi \equiv x/\lambda z \cong \sin\theta_m/\lambda$. The intensity of the m -th diffraction order is given by $I_m = |O(\theta_m)|^2$. Therefore the diffraction efficiency of the m -th order can be expressed as

$$\eta_m \equiv \frac{I_m}{I_{\text{incident}}} = 4 \left[\frac{\sin(m\pi(w/d))}{m\pi} \right]^2 \cos^2(\phi_0 + \pi m/2). \quad (10.4d)$$

From Eq. (10.4d), for a 50% duty cycle, i.e., $w/d = 2$, the maximum diffraction efficiency for two equal beams occurs when $\phi_0 = 90^\circ$, which produces $I_0 = 0$, $I_{\pm 1} = 4/\pi^2$, and $\eta = 81\%$. For three equal beams to occur, the phase angle needs to be $\phi_0 = 57.52^\circ$, which results in $I_0 = I_{\pm 1}$ and $\eta = 86.5\%$. For larger numbers of diffracted beams, it is not possible to obtain equal amplitude solutions with Eq. (10.4d) for a fixed w/d because there are not enough free variables to satisfy the equations needed for equal intensities.

However, by using different groove widths within an individual grating period, multiple beams can be formed with equal intensities. One design approach to binary phase gratings is adapted from the technique of Lee [10.27, 10.28]. For this approach, parameters including transition points x_1, x_2, \dots, x_q and the phase angle ϕ_0 are determined and optimized so that N central diffraction orders will not only have the maximum diffraction efficiency but also have equal intensities (Fig. 10.2b).

The intensity distribution of N beams in the far field is calculated by taking the Fourier transform of each slit of the grating and summing the diffracted intensity. Here, a slit is defined as the interval between two adjacent transition points, and uniform illumination on the grating is assumed.

The diffracted complex amplitude in the far field from a single slit of the grating is described by

$$A_n(\xi) = \mathcal{F} \{ \text{rect}(x/w_n) e^{-i\phi_0} \} . \quad (10.5)$$

where $\xi \equiv x/\lambda z \cong \sin\theta_m/\lambda$; K is a normalization factor; n is the n -th slit, w_n is the width of the n -th slit defined as $w_n = x_{i+1} - x_i$; θ_m is the diffraction angle of the grating measured in radians; and ϕ_0 is the phase term of the n -th slit. The ϕ_0 in Fig. 10.2b is the phase change relative to a substrate as experienced by light passing through its thickness, and is related to the refractive index of the phase layer n and phase layer thickness t by $\phi_0 = (2\pi/\lambda) \cdot (n-1)t$.

Referring to Eq. (10.5), the total complex amplitude diffracted by the illuminated grating is the summation of each individual diffracted amplitude from each illuminated slit:

$$A(\xi) = \sum_{k=0}^M \sum_{n=1}^q K w_n \text{sinc}(w_n \theta_m / \lambda) [(\cos\phi_{nk} + i \sin\phi_{nk})] . \quad (10.6)$$

where ξ is defined as before; M is the total number of illuminated grating periods; q is the number of slits within a grating period; K is again the normalization factor; ϕ_{nm} is the phase term due to the n -th slit of the m -th illuminated grating period and can be expressed as

$$\phi_{nm} = (2\pi/\lambda)(md + x_n) \sin\theta_m + \phi_0 . \quad (10.7)$$

where m is the m -th grating period illuminated by the incident beam; d is, again, the grating period; x_n is the center location of n th slit.

The grating period d and the number of grating periods M required to be illuminated are determined by the application. One example is illustrated in Fig. 10.3, which depicts an optical system for generating a set of focused spots to switch an array of optical logic gates. The dimension of each GaAs etalon in the array is assumed to be approximately $9 \mu\text{m} \times 9 \mu\text{m}$ square. The separation Δh between two gates has to be at least two times larger than the gate's dimension to prevent significant crosstalk by means of diffusion. Referring to Fig. 10.3, for a normally incident beam ($\theta_i = 0$), the grating equation is

$$d(\sin\theta_m - \sin\theta_i) = d\sin\theta_m = m\lambda \quad m = 0, \pm 1, \pm 2, \dots, \pm N/2, \dots$$

or

$$\frac{dh_m}{\sqrt{f^2 + h_m^2}} = m\lambda \quad (10.8)$$

where m is the diffraction order; θ_m is the diffraction angle of the m -th diffraction order; N is the number of equal beams; h_m is the distance between the center of the m -th beam and the optical axis, which can be expressed from Eq. (10.8) as

$$h_m = \frac{f}{\sqrt{(d/m\lambda)^2 - 1}} \quad (10.9)$$

From Eq. (10.9), the grating period must be greater than the diffraction order times the illuminating wavelength.

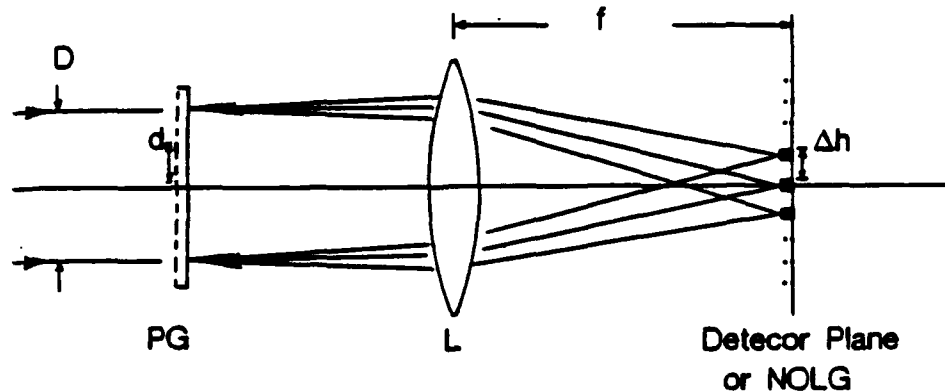


Figure 10.3 Scheme of optical system to generate an array of focused spots for the multiple switching on optical logic gates.

In general, since the focal length f of a lens used in this optical system is much greater than the separation of two beams Δh , for small angle approximation (if $\theta \leq 15^\circ$, then $\sin\theta \approx \theta$ with relative error $\leq 1\%$) we have

$$\Delta h \approx \lambda f / d \quad (10.10)$$

For diffraction-limited resolution and non-overlapping diffraction pattern, one needs

$$2 \times 2.44\lambda(f/D) \leq \Delta h, \quad (10.11)$$

where D is the diameter of the incident beam. Substituting Eq. (10.10) into Eq. (10.11), we have $M = D/d \geq 4.88$, that is, the diameter of the incident beam should at least illuminate five grating periods to prevent crosstalk during the switching operation of a GaAs-etalon array.

The diffracted intensity of each order I_1, I_2, \dots, I_N can be obtained by multiplying the amplitude given in Eq. (10.6) by its complex conjugate. Optimization is necessary in order to find the maximum diffraction efficiency $\eta(w_1, w_2, \dots, w_q, \phi_0)$ under the constraint of $I_1 = I_2 = \dots = I_N$. Numerical methods such as Lagrangian techniques can be applied in the optimization [10.34]. The number of transition points is determined by the existence requirement of the solution, namely, the number of parameters in η should equal the number of constraints set by the requirement of $I_1 = I_2 = \dots = I_N$ and a maximum η . When N is even, $N = 2p$,

$$\eta = \left[2 \sum_{k=1}^p I_k \right] / I_{\text{incident}}, \text{ and when } N \text{ is odd, } N = 2p+1, \eta = \left[2 \sum_{(k=1)}^p I_k + I_0 \right] / I_{\text{incident}}.$$

The required number of transition points q for odd and even values of p can be given by [10.29]

$$\begin{aligned} q &= p+1 \text{ for odd } p \\ q &= p+2 \text{ for even } p. \end{aligned}$$

The design procedure was programmed and run on a personal computer. Parameters $w_1, w_2, \dots, w_q, \phi_0$ were varied until a maximum overall diffraction efficiency with equal amplitude in central diffraction orders was obtained. Figure 10.4 shows the phase profile of a nine-beam grating and the resultant diffracted intensity pattern. Application of nine- and two-beam binary phase gratings are demonstrated in the three-spot pattern recognition experiment in Section 13. The following section discusses a proposal for fabricating binary phase gratings.

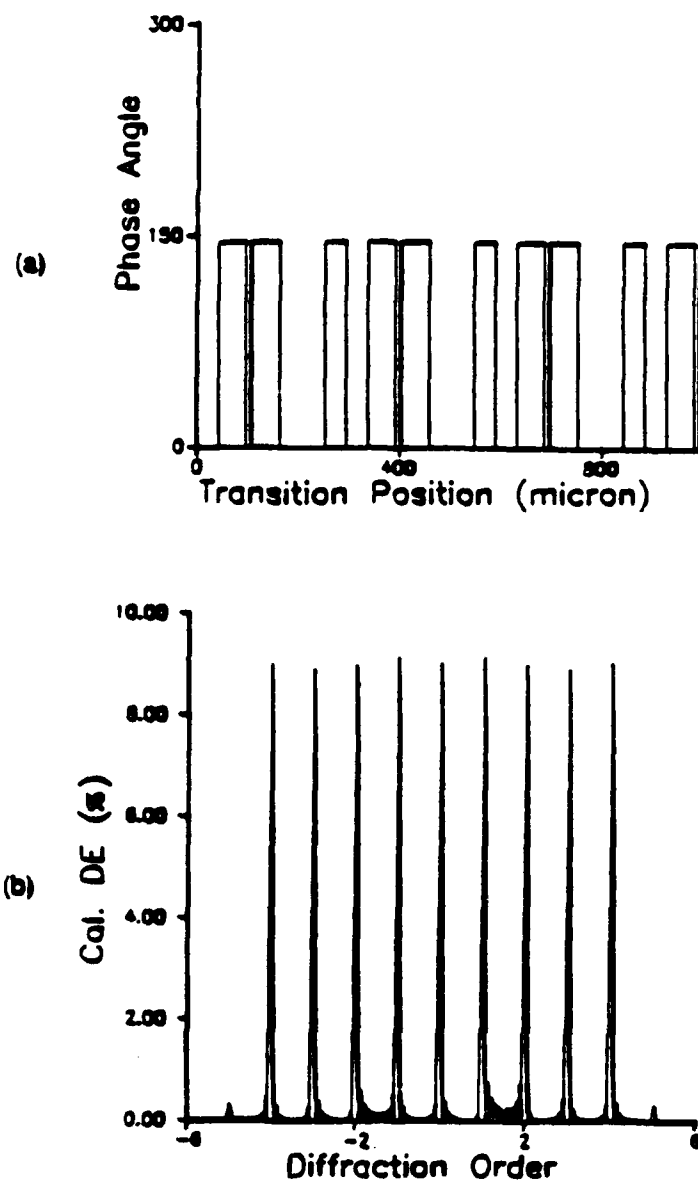


Figure 10.4 Design of a binary phase grating for generating nine equal beams; (a) phase profile, (b) intensity distribution.

Grating Fabrication

Binary phase gratings can be fabricated using IC lithographic techniques. The fabrication process involves substrate preparation, phase-layer coating, photomasks generation and contact copying, plasma etching, and testing. Other methods employing direct encoding and etching on a glass substrate have been demonstrated [10.27, 10.28]. Control of the etching depth is a critical parameter in these techniques.

A typical wafer of single-crystal sapphire (Al_2O_3) with c-axis parallel to the beam direction can be used as a substrate ($n \cong 0.59$). Surface quality of 60-30 (scratch-dig) and $\lambda/4$ flatness is generally required for the substrate.

A transparent dielectric material is used for the groove material, for example, silicon dioxide for infrared to visible. Moreover, gratings made of silicon dioxide are lightweight and rugged, with high damage threshold, and are relatively insensitive to changes in temperature and humidity. A SiO_2 layer can be coated on a sapphire substrate using either chemical vapor deposition (CVD) or electron-gun sputtering techniques. A thin, quarter-wave SiO_2 layer can also be deposited on the back of the substrate as an antireflection coating to further reduce the Fresnel loss of 5% to 2%. The coated phase layer should be several hundred angstrom thicker than the design thickness, allowing deviation due to the uncertainty of refractive indices of a thin film material. An ellipsometer can be used to measure the refractive index of SiO_2 film at a certain wavelength; refractive indices at other wavelengths can be estimated through the interpolation method. Typical refractive indices of SiO_2 used for the designs discussed in the previous section were $n \cong 1.465$ for $\lambda = 632.8$ nm, and $n \cong 1.481$ for $\lambda = 514.5$ nm, respectively.

A photomask is used to transfer the phase profile to the SiO_2 layer. The mask consists of a series of chromium stripes laid on a flat glass substrate. The width and the location of each stripe, that is, the phase transition points, are determined by the previously mentioned calculations.

Tolerance analysis was performed to determine the effect of positioning inaccuracies of the phase transition points. These inaccuracies may be caused by the limited resolution of common plotting devices. Each phase transition requires more accurate positioning when the number of beams increases and/or when the diffraction angle becomes large. Either performing a simulation in the design program or referring to Eq. (10.8) and the relationship between N and p previously mentioned

can indicate these requirements. These requirements can be met by using an electron-beam writing facility which offers high resolution and can be employed to encode a pattern for the photomask of a binary phase grating [10.19].

The pattern is copied in the following way. First, a layer of photoresist is deposited on the SiO_2 layer after an appropriate thickness of SiO_2 has been coated on the substrate wafer. Then, the photomask is brought in close contact with the wafer and is illuminated. The stripes of the photomask that are clear allow the light to pass through and expose the photoresist. If the photoresist type is negative, these illuminated stripes of the photoresist are then chemically washed away to expose the SiO_2 . The exposed SiO_2 is then plasma etched down to the substrate and the rest of the photoresist is then chemically removed. At this point, one may have preliminary phase structure.

Two steps in the grating fabrication require rigorous testing: stripe-width measurements of a photomask and performance evaluation of etched phase gratings. The former step can be done by using an optical comparator or a high power microscope or even by a SEM for fine structures. Special attention is needed to ensure that a measurement device can resolve the narrowest-width stripe. Performance evaluation of an etched phase grating involves iterative interactions between the etching process and the intensity measurement to ensure that a satisfactory intensity distribution is obtained. The etching process is irreversible. Each removal step takes off thickness of several tens of angstroms and is followed by intensity measurements of the diffracted beams. Intensity measurements should be made at several locations over the surface of the grating-filled wafer to examine the fabrication quality. After these testing steps, deviations from the expected diffracted intensities can be removed by iterated etching. Using a tolerance analysis similar to the previously mentioned (except for phase thickness instead of phase transition points), one can prevent over-etching; however the success of fabricating a phase grating heavily depends on the interactions between these two steps and practical knowledge of the etch process.

10.3 Thick Phase Hologram

In this section, the theory of a thick planar phase hologram is reviewed. Then, thin and thick holograms are compared, and the situation in which all holograms made were thick enough to have Bragg-regime diffraction is discussed. Finally, theoretical predictions and experimental results from holograms found in dichromated gelatin (DCG) are compared.

Review of Coupled Wave Analysis For a Single Grating

To determine the diffraction efficiency from a thick phase hologram, we must solve the wave equation for the fields propagating through a medium with a sinusoidal index variation. If a phase grating is sufficiently thick and of sufficiently high spatial frequency, then for reconstruction near the Bragg angle, only the incident wave and a single diffracted wave are observed. Under these conditions, a coupled wave theory developed by Kogelnik [10.35] can be used to derive a closed-form analytic solution to the diffraction efficiency problem.

In Fig. 10.5, all the waves are polarized perpendicular to the plane of incidence (s-polarized waves). Only the reconstruction beam R and the diffracted signal beam S are shown. Other diffracted orders strongly violating the Bragg condition are not observed experimentally, and are neglected in the analysis. The fringe planes are assumed to extend infinitely in the \hat{y} direction and are slanted with respect to the medium boundaries at an angle ϕ . The grating vector \vec{K} is normal to the fringe planes and has the magnitude $K = 2\pi/d$ where d is the grating period. The incident angle in air is θ_i and is related to the angle inside the medium θ_r by Snell's law, $\sin\theta_i = n_0 \sin\theta_r$, where n_0 is the average refractive index of the medium. The recording material is assumed to be lossless and nonmagnetic ($\mu = 1$); therefore discussion hereafter will focus on phase gratings without bulk, without modulated absorption, and without magnetic properties.

Wave propagation in the grating is described by the scalar wave equation

$$\nabla^2 E(x,z,\omega) + k^2(x,z) E(x,z,\omega) = 0 \quad , \quad (10.12)$$

where $E(x,z,\omega)$ is the complex electric field and is assumed to be independent of \hat{y} , that is, the fringe planes are perpendicular to the plane of incidence; ω is the oscillation radial frequency; $k(x,z)$ is spatially modulated and related to the relative dielectric constant $\epsilon(x,z)$ by

$$k^2(x,z) = \frac{\omega^2}{c^2} \epsilon(x,z) \quad . \quad (10.13)$$

Since the interference pattern produced by two plane waves has a sinusoidal intensity distribution, a linear recording and developing process produces a modulation in the region $0 \leq z \leq T$, given by

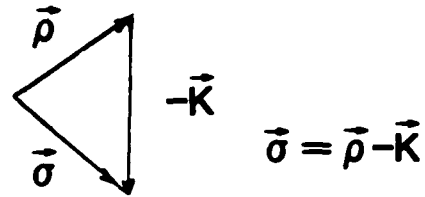
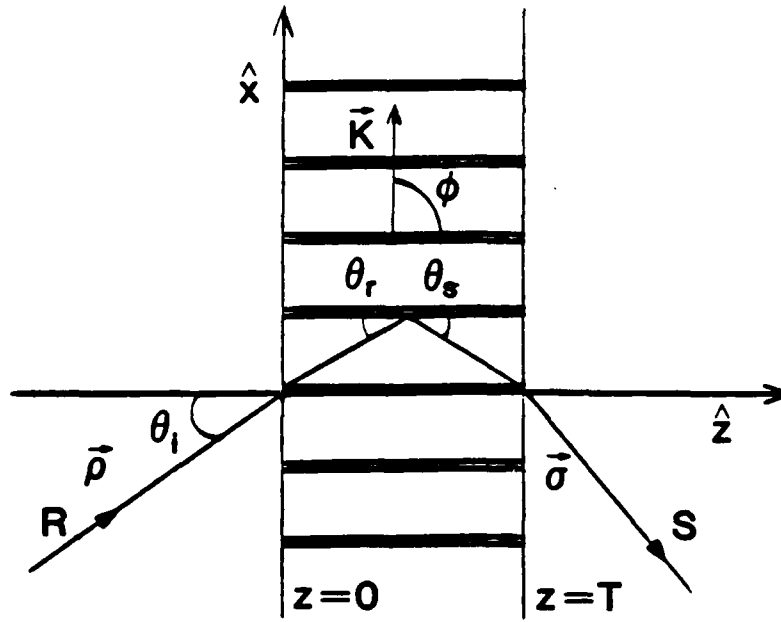


Figure 10.5 Geometric structure of a thick hologram formed by two plane waves, both having s polarization.

$$\epsilon = \epsilon_0 + \Delta\epsilon \cos(\vec{K} \cdot \vec{r}), \quad (10.14)$$

where ϵ_0 is the average dielectric constant of the material; $\Delta\epsilon$ is the amplitude of modulated permittivity; and $\epsilon_0 \gg \Delta\epsilon$. The grating vector \vec{K} and position vector \vec{r} have the form of

$$\vec{K} = K \begin{bmatrix} \sin \phi \\ 0 \\ \cos \phi \end{bmatrix} \quad \text{and} \quad \vec{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (10.15)$$

respectively. Substituting Eq. (10.14) into Eq. (10.13), one has

$$k^2 = \beta^2 + 2\kappa\beta(e^{i\vec{K}\cdot\vec{r}} + e^{-i\vec{K}\cdot\vec{r}}) , \quad (10.16)$$

where the average propagation constant β and the coupling coefficient κ are defined as

$$\beta \equiv 2\pi\epsilon_0^{1/2}/\lambda = 2\pi n_0/\lambda \quad (10.17)$$

and

$$\kappa \equiv \pi\Delta\epsilon/2\lambda\epsilon_0^{1/2} = \pi\Delta n/\lambda . \quad (10.18)$$

If the reconstruction wave is s polarized and incident near or at the Bragg angle, the total electric field in the grating is the superposition of two plane waves (R and S) with propagation vectors $\vec{\rho}$ and $\vec{\sigma}$:

$$E = R(z)e^{-i\vec{\rho}\cdot\vec{r}} + S(z)e^{-i\vec{\sigma}\cdot\vec{r}} . \quad (10.19)$$

The grating vector \vec{K} provides the coupling mechanism between the incident wave R and the diffracted signal wave S. The energy interchange occurs as the wave propagates through the grating (\hat{z} direction). The amplitudes of the waves do not depend on x since the grating is invariant in this direction. The incident wave R is diffracted by the grating \vec{K} , generating the wave S. The Bragg (K-vector closure) condition yields two equations

$$\begin{aligned} \vec{\sigma} &= \beta \begin{bmatrix} -\sin\theta_s \\ 0 \\ \cos\theta_s \end{bmatrix} \\ &= \vec{\rho} - \vec{K} = \begin{bmatrix} \beta\sin\theta_r - K\sin\phi \\ 0 \\ \beta\cos\theta_r - K\cos\phi \end{bmatrix} \end{aligned} \quad (10.20)$$

(all θ_r, ϕ measured inside the medium) .

Equivalently, one has

$$(\sin\theta_r + \sin\theta_s) = \lambda/n_0 d \sin\phi \quad (10.21a)$$

and

$$\beta(\cos\theta_r - \cos\theta_s) = \lambda/n_0 d \cos\phi . \quad (10.21b)$$

Note that when the reconstruction wave R is incident at the Bragg angle, both the reflection law $\theta_r = \theta_s$ (10.21b), and the diffraction condition $2\sin\theta_r = \lambda/n_0d$ (10.21a) hold. Interference between the wave R and the wave S produces fringes that are composed of many planes of constant index-of-refraction across the interference area. These planes are oriented in such a way that the reconstruction wave R is reflected into the signal wave S. Since each change in the refractive index produces some reflection on reconstruction, the wave R transverses many of these index planes; therefore successively reflected waves are exactly in phase with each other, and hence add coherently to produce the diffracted wave S. Other orders do not have this mirror geometry and are dephased by subsequent grating planes; therefore for non-Bragg angle the successively reflected plane waves are slightly out of phase with each other and do not produce a strong, coherent superposition. This "dephasing" of the successively diffracted plane waves is employed by the coupled wave theory to predict the reduced diffraction efficiency (DE) for non-Bragg angle illumination.

Substituting Eq. (10.16) and Eq. (10.19) into Eq. (10.12) and applying the Bragg condition to collect terms with a common propagation direction, dropping terms $\exp(-i(\vec{\rho}+\vec{K})\cdot\vec{r})$ and $\exp(-i(\vec{\sigma}-\vec{K})\cdot\vec{r})$, one has the following coupled differential equations describing the energy flow between R and S:

$$\begin{aligned} R'' - i2\rho_z R' + 2\kappa\beta S &= 0 \\ S'' - i2\sigma_z S' + 2\beta\Gamma S + 2\kappa\beta R &= 0 \end{aligned} \quad (10.22)$$

Apparently, the reconstruction wave R and the diffracted wave S are coupled with each other through the coupling coefficient κ . Here, one introduces Γ as a parameter to indicate deviations from the Bragg condition and is expressed as

$$\begin{aligned} \Gamma &= \frac{\beta^2 - \sigma^2}{2\beta} \\ &= K\cos(\phi - \theta_r) - K^2\lambda/4\pi n_0 \\ &= \Delta\theta_r K\sin(\phi - \theta_{r,b}) - \Delta\lambda K^2/4\pi n_0 \end{aligned} \quad (10.23)$$

where $\theta_r = \theta_{r,b} + \Delta\theta_r$ and $\lambda = \lambda_b + \Delta\lambda$; $\theta_{r,b}$, λ_b are the Bragg angle and wavelength respectively while $\Delta\theta_r$ and $\Delta\lambda$ are the amount of deviation from these values.

If the slowly varying envelope approximation is made, R'' and S'' will be negligible as compared with $\rho_z R'$ and $\sigma_z S'$, respectively. Under these conditions the coupled wave equations become

$$\begin{aligned} C_R R' &= -i\kappa S \\ C_S S' + i\Gamma S &= -i\kappa R \end{aligned} \quad (10.24)$$

where $C_R \equiv \rho_z/\beta = \cos\theta_r$ and $C_S \equiv \sigma_z/\beta = (\beta\cos\theta_r - \kappa\cos\phi)/\beta$ are its obliquity factors for the wave R and the wave S.

Solutions for R and S are obtained by applying the boundary conditions $R(z = 0) = 1$, $S(z = 0) = 0$. For an unslanted transmission grating ($\phi = \pi/2$), diffraction efficiency at the Bragg angle is

$$\begin{aligned} DE &= \left| \frac{C_S}{C_R} \right| S(z = T) S^*(z = T) \\ &= \sin^2 \gamma \end{aligned} \quad (10.25)$$

where γ is the angular modulation or grating strength parameter, and is defined as

$$\gamma = \frac{\pi \cdot \Delta n \cdot T}{\lambda \cdot \cos\theta_{r,b}} \quad (10.26)$$

= $(\pi/\lambda) \cdot \Delta n \cdot (\text{geometrical optical path length experienced by the beam transversing hologram})$.

For the non-Bragg-angle condition, diffraction efficiency will be

$$DE = \frac{\sin^2 \left[\gamma \sqrt{1 + (\xi/\gamma)^2} \right]}{1 + (\xi/\gamma)^2} \quad (10.27)$$

where ξ is the dephasing term

$$\begin{aligned} \xi &= -\Gamma T / 2C_S \\ &= \Delta\theta_r \pi T / d \text{ (angular deviation for } \lambda_b) \\ &= -(\Delta\lambda/\lambda) \cdot T \beta \tan\theta_{r,b} \sin\theta_{r,b} \text{ (wavelength deviation for } \theta_{r,b}) \end{aligned} \quad (10.28)$$

Referring to Eq. (10.28), therefore, the peak efficiency and spectral bandwidth are controlled by index modulation, film thickness and angle of incidence of the recording beam. A rule of thumb for angular selectivity can be expressed by $\Delta\theta_r \cong d/T = 1/(\text{total no. of fringes crossed in propagation through the emulsion})$, which primarily depends on the emulsion thickness and is independent of index modulation.

"Thin" and "Thick" Grating Criteria

The characteristics derived above are based on the assumption that a grating is thick enough to satisfy the Bragg condition. However, one might ask how thick a grating must be to ensure the occurrence of the properties previously mentioned, e.g., possibility of obtaining unity diffraction efficiency, and strong angular and wavelength selectivity.

Traditionally, a thin grating is defined as a grating that can produce Raman-Nath diffraction (many diffraction orders) while a thick grating can produce Bragg diffraction (single diffraction order). It is better to have some criteria to distinguish between these two regimes. According to the analysis by Gaylord et al. [10.24], and a more rigorous analysis by Jaaskelainen et al. [10.36], one can use the parameters Q' , γ , ρ to distinguish these two regimes for unslanted phase gratings, where Q' , ρ are defined respectively as:

$$\begin{aligned} Q' &\equiv (2\pi\lambda T)/(n_0 d^2 \cos\theta_r) \\ \rho &\equiv Q'/2\gamma = \lambda^2/(d^2 n_0 \Delta n) \end{aligned} \quad (10.29)$$

Figure 10.6 shows Raman-Nath, intermediate, and Bragg diffraction regimes and their corresponding boundaries:

- (1) The Raman-Nath diffraction regimes (for thin gratings) occurs when

$$Q'\gamma \leq 1 \text{ and } \tan^2\theta_r \ll Q'. \quad (10.30)$$

- (2) The Bragg diffraction regimes (for thick gratings) occurs when

$$\rho = Q'/2\gamma \geq 1. \quad (10.31)$$

Jaaskelainen et al. [10.36] also showed that as the angle of incidence grows the Raman-Nath regime will shrink and move toward the boundary where $Q'\gamma = 1$; moreover when the incident angle is greater than 15 degrees, the Raman-Nath regime will disappear. It is interesting to note that according to the above criterion the definition of a thick grating is independent of grating thickness. This indicates that Bragg diffraction will occur for any value of γ if Q' is sufficiently large. An alternative expression of Eq. (10.31), expressed as follows, can be derived which shows the angle of incidence needed to obtain a thick grating, i.e.,

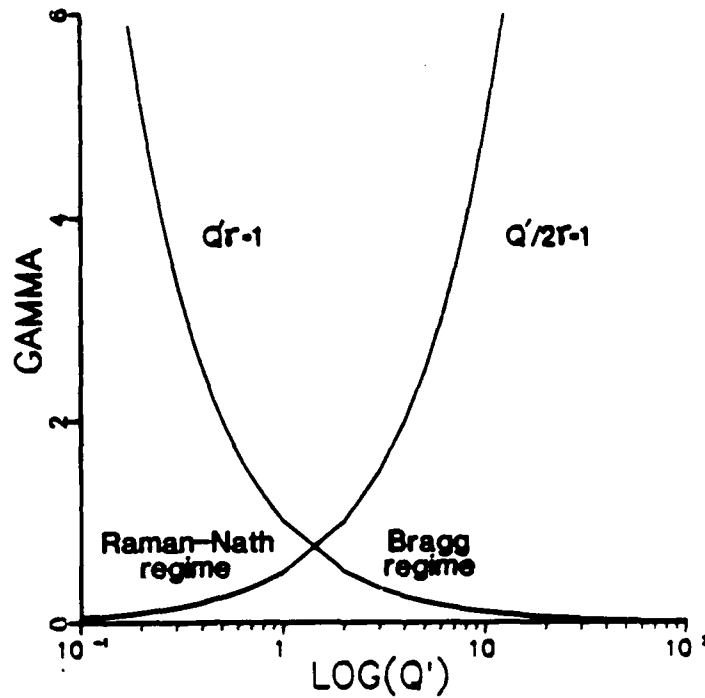


Figure 10.6 Regimes of Raman-Nath, intermediate, and Bragg diffraction.

$$\sin\theta_i \geq \sqrt{2.5n_0\Delta n} \quad (10.32)$$

For Kodak 649F derived DCG recording material [10.37], $n_0 = 1.54$ to 1.55 and Δn can be varied in the range 0.01 to 0.08 with exposure. Therefore, the interbeam angle θ_i needs to be greater than 22 to 68 degrees to obtain a thick grating.

For sinusoidal phase gratings in the Raman-Nath regime, diffraction efficiency of the i -th diffracted beam can be expressed as [10.24]

$$DE_i = J_i^2(2\gamma) \quad (10.33)$$

where J_i is the integer-order Bessel function of the first kind, and γ is given by Eq. (10.26). Note that the maximum diffraction efficiency in the Raman-Nath regime is independent of Q' and equals 33.8% when $\gamma = 0.1425$. The diffraction efficiency of sinusoidal phase gratings can be 100% in the Bragg regime when $\gamma = \pi/2$ in Eq. (10.25).

10.4 Experiments on Planar Thick Holograms

Thick phase gratings using DCG emulsion can produce high diffraction efficiency (Appendix B). However, variations in the diffraction efficiency from DCG emulsion exist because of the complicated dependence of their characteristic properties on parameters like temperature, humidity, chemical processing, and environment. To achieve diffraction efficiency as high as possible from a certain chemical processing, one needs to explore every possible effect of these different parameters. Although some conditions such as temperature and humidity are not easily controlled in our laboratory, an optimized process has been achieved with consistent results.

In the following section, experimental results obtained from planar thick holograms are described and are compared with theory. Results from high-exposure recording and possible observations of these holograms are then discussed.

The DCG plates used in this study were prepared with the procedures outlined in Appendix B. The various experimental aspects of this study are given here.

Thick holographic gratings in DCG were recorded by mixing two mutually coherent plane waves with an interbeam angle of about 36° , producing unslanted transmission gratings. The recording parameters for these holograms are $\lambda = 0.5145 \mu\text{m}$; $T = 11.2 \mu\text{m}$; $n_0 = 1.545$; and $d = 0.83 \mu\text{m}$, which give a Q' factor, from Eq. (10.29), of 35. The Q' factor has been kept relatively constant throughout the study. For Bragg-regime diffraction Eq.(10.31) Δn should be at least less than 0.025. The diffraction-efficiency measurements were made using the Hamamatsu S1226 silicon photodetector and Keithley 485 Autoranging picometer. The diffraction efficiency is calculated by using the formula

$$DE = \frac{I_1}{I_0} \quad (10.34)$$

where I_1 is the diffracted intensity of the first order and I_0 is the intensity of the incident beam. Reflection losses are not included.

A number of gratings have been recorded in DCG by exposing them to green light from a 5-W Model 2020 Spectra Physics argon ion laser. The laser was operated with an etalon to provide single-mode longitudinal operation. This was verified by measurements with a scanning Fabry-Perot. Figure 10.7 shows the curve of diffraction efficiency (DE) versus exposure for planar gratings read at the Bragg angle when $E < 330 \text{ mJ/cm}^2$, and at off-Bragg angle when $E > 330 \text{ mJ/cm}^2$ for the maximum diffraction efficiency. Non-Bragg-angle efficiency measurements in the latter case were made because the maximum efficiency no longer occurred at the

Bragg angle. The index modulation was suspected to be saturated and will be discussed in detail later. In contrast, a theoretical diffraction efficiency versus Δn curve for a single planar grating described by Eq. (10.25) is plotted in Fig. 10.8 by taking previous data of λ , T , n_0 , and d . Several significant differences between experimental results and theory are summarized and discussed as follows.

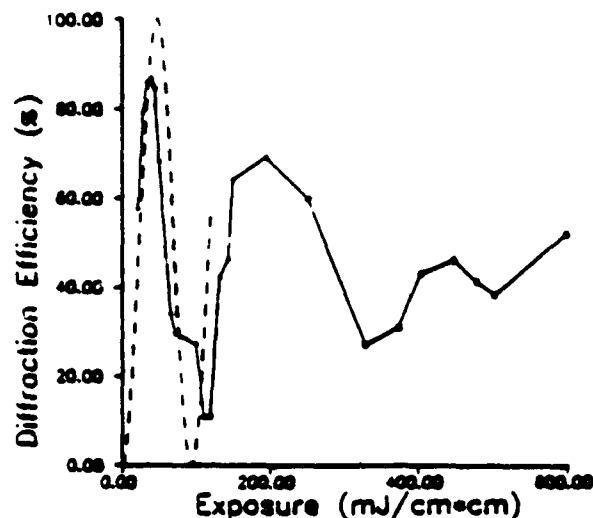


Figure 10.7 DE versus exposure for transmissive, planar gratings in DCG material. The dashed line is the calculated curve for linear exposure region. Experimental data shown in squares were taken for the maximum DE at non-Bragg angle.

First, the maximum diffraction efficiency of experimental data is about 87%, without considering losses resulting from reflection, reflection-diffraction, and high-order diffractions as depicted in Fig. 10.9. Reflection-diffraction loss can be distinguished from reflection loss in the sense that reflection-diffraction loss is caused by reflection from the last surface, which produces a weak beam propagating in the opposite direction of S , thus reconstructing a weak retro-reflected beam when the reconstruction beam passes through the recorded grating. In contrast, reflection loss is the total loss contributed by each reflection when the reconstruction beam passes each interface. Therefore, reflection-diffraction loss can be eliminated only when the last surface is anti-reflection coated, as evidenced by the observation that this beam disappears if an oil-immersed black cloth is affixed to the last surface.

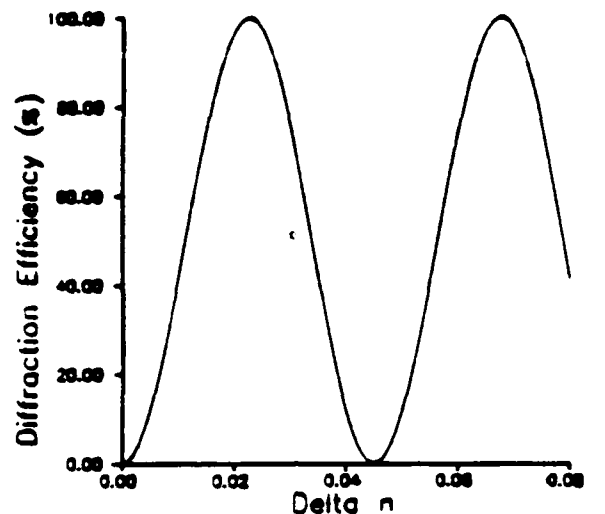


Figure 10.8 Calculated DE versus Δn .

These reflection losses can produce a 6% to 10% reduction in diffraction efficiency. High-order diffraction beams (only $m = \pm 3$ without $m = \pm 2$!) were observed from both sides of a grating, indicating some "weak" thin grating effect despite the large Q' factor. The loss attributable to these high-order diffraction beams comprises about 1% of the total diffraction efficiency; however this loss can be eliminated or suppressed by increasing the incident angle and/or beam ratio (R/S) [10.38]. Taking all these losses into account, one can reach at least 94% diffraction efficiency for the transmission holographic phase gratings found in this medium without index matching. It is expected that diffraction efficiency can be improved by recording the hologram in a liquid gate.

The diffraction efficiency curve in Fig. 10.7 is similar to that predicted by Fig. 10.8 at low exposures (such as $E \leq 120 \text{ mJ/cm}^2$). The maximum diffraction efficiency obtained when $\Delta n = 0.022$ in Fig. 10.8 is equivalent to the experimental exposure of ≈ 35 to 40 mJ/cm^2 in Fig. 10.7.

The agreement between theory and experiment at high exposure is not very good, indicating that the recording process is no longer linear and the "linear recording" assumption in the coupled-wave theory has been violated. The following procedures were taken to determine the linearity range of Δn versus exposure.

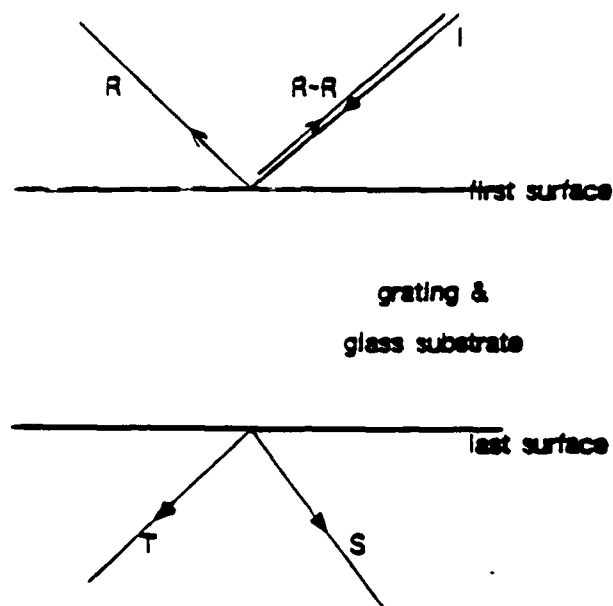


Figure 10.9. Loss observed when a holographic grating is reconstructed by an incident beam *I*: *R-R*, reflection-diffraction; *R*, reflection; *S*, signal; *T*, transmission.

Figure 10.7 shows data from 20 different plates and represents the reproducibility of the process. Using these data in conjunction with Eq. (10.25), the relation between Δn and E can be determined. Figure 10.10 shows derived values for Δn versus E for an exposure $E \leq 120 \text{ mJ/cm}^2$ (in squares). The solid line approximates a linear relation between Δn and E according to the relation,

$$\Delta n = 0.000385E (+0.006354) \quad (10.35)$$

This relation and Eq. (10.25) can be used to calculate the diffraction efficiency as a function of the exposure (see the dashed line in Fig. 10.7).

The experimentally measured data of the diffraction efficiency reaches a maximum value of 90%. As Δn increases further, the efficiency decreases approximately a value of 11% near $E = 115 \text{ mJ/cm}^2$ and then starts to increase again. This indicates that a decrease in an index modulation with increased exposure occurs even under optimum exposure conditions. This effect also occurs for materials other than DCG, and can be seen in silver-based material [10.39], probably because of saturation of the index modulation [10.40]. This effect could also be the result of a

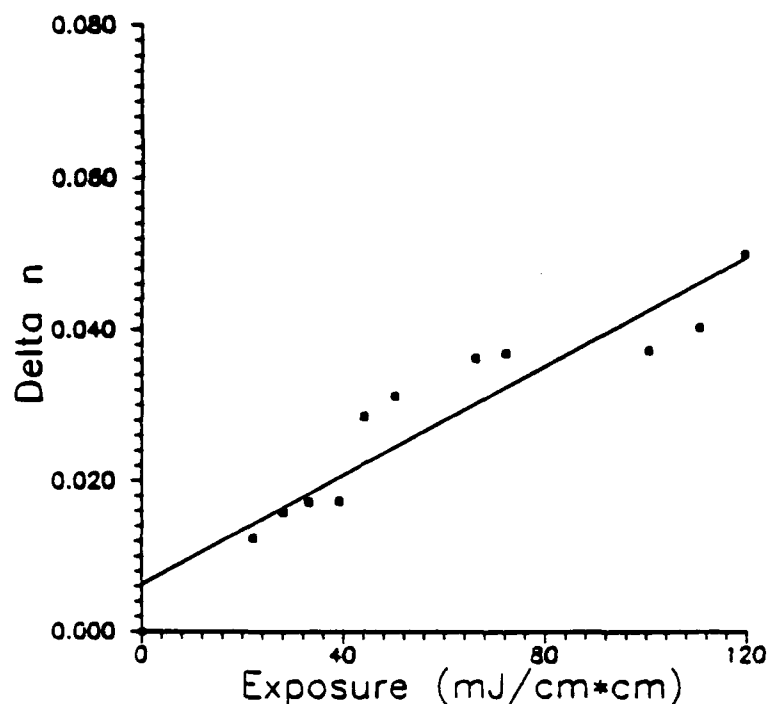


Figure 10.10 Calculated Δn versus experimental exposures.

violation of the assumptions proposed in the coupled wave analysis, i.e., the uniform index modulation with depth (in the z direction). The sensitized film is strongly absorbing and attenuates light in the z direction in a fairly complicated manner [10.40]. We conclude that a linear relationship between Δn and exposure cannot always be assumed over a large exposure range. Therefore a diffraction efficiency curve does not behave sinusoidally for high exposures.

Angular and wavelength selectivities are important characteristics of gratings which provide information about emulsion thickness and index modulation Δn . Figure 10.11 shows the predicted (dashed line) and measured angular selectivity of a planar grating with total exposure of approximately 35 mJ/cm^2 (Fig. 10.7) so that $\gamma \cong \pi/2$. After taking losses into account, the experimental results show fairly good agreement with theory. Angles were measured with a NRC Model 481 Rotary Stage which has resolution of 30 arc sec. From the width of the curve, the film thickness is determined to be $11.2 \text{ }\mu\text{m}$ by Eq. (10.28). Actual thickness measurement over several locations on a developed film was made by a standard DEKTAK stylus (0.0005-in. radius) with scan speed 0.1 cm/min. , which showed that the emulsion was about $11.5 \text{ }\mu\text{m}$ thick, in good agreement with the calculated one. Note that near $\pm 40^\circ$,

where the calculated DE = 0, the measured DE \neq 0. The discrepancy could result from the index modulation, Δn , being nonuniform with depth, z , as previously mentioned above.

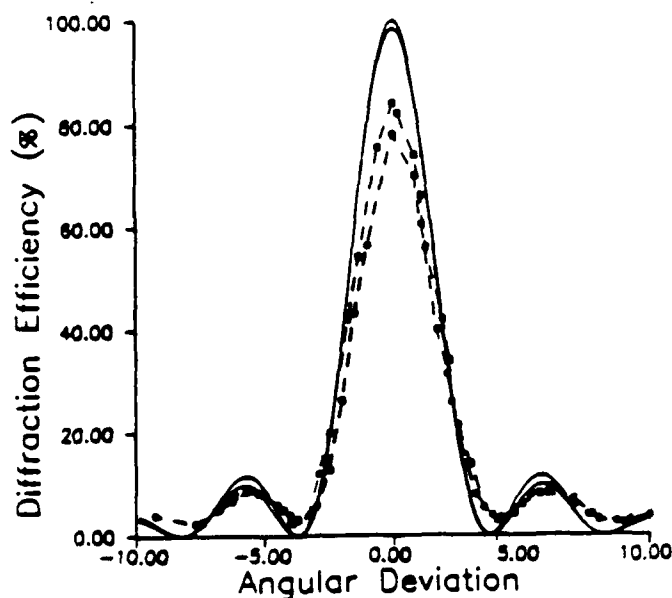


Figure 10.11 Angular selectivity of a thick phase grating of $E \approx 35 \text{ mJ/cm}^2$. Solid lines are calculated values for s (higher one) and p polarization. Experimental data are shown in squares for s polarization and in circles for p polarization.

Figure 10.12 shows the measured angular selectivity of a planar grating with an exposure $E = 330 \text{ mJ/cm}^2$, where the solid line is the closest simulation using parameters $T = 11.2 \text{ }\mu\text{m}$ and $\Delta nT = 0.94 \text{ }\mu\text{m}$, which leads to $\gamma \approx 1.9\pi$. The difference between experimental and calculated data exists because the recording process is no longer linear. Moreover, additional diffracted orders began to emerge, which along with the nonlinear recording violate the assumptions of the coupled-wave analysis. Both simulated and experimental data taken at high exposure are quite different from data at low exposures. In addition, the diffraction efficiency at the Bragg angle is lower than the diffraction efficiency for high exposures. This is assumed to result from over modulation of the refractive index.

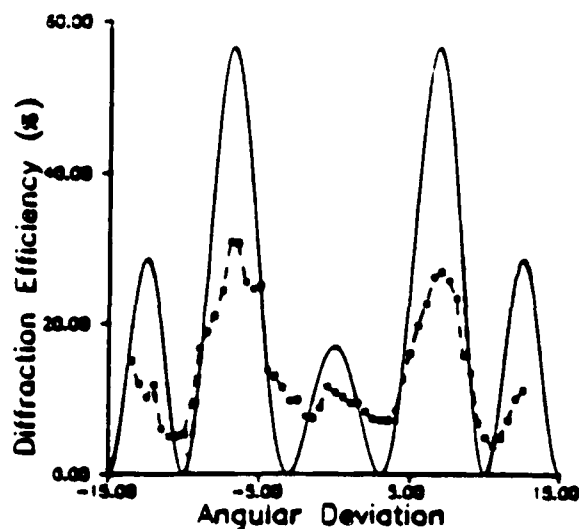


Figure 10.12 Angular selectivity of a thick phase grating of $E \cong 330 \text{ mJ/cm}^2$. The solid line is for calculated values with $\Delta nT = 0.94 \text{ } \mu\text{m}$ and the dashed line for experimental data.

The spectral selectivity can be measured by recording a planar grating at one wavelength and measuring the diffraction efficiency at another wavelength. Figure 10.13 shows that the diffraction efficiency varies over eight discrete wavelengths, ranging from near IR to visible (in squares) and the predicted curve is also shown in the dashed line. The holographic grating was recorded at an exposure of 45 mJ/cm^2 and $\lambda = 0.5145 \text{ } \mu\text{m}$. Six prominent lines in an argon-ion laser, one He-Ne laser, and a laser diode of Sharp Type LTO22MCO ($\lambda = 0.786 \text{ } \mu\text{m}$) were used for this readout. It can be observed that the diffraction efficiency has a maximum value at the recording wavelength and decreases when the wavelength changes from this value. Note that the experimental data for the long-wavelength range, i.e., $\lambda = 0.6328$ and $0.786 \text{ } \mu\text{m}$, were taken for the maximum diffraction efficiency instead of the diffraction efficiency at the Bragg angle because of the experimental difficulties stated as follows. If separate lasers other than the one used in the recording were used to perform spectral selectivity measurement, then (i) it was hard to locate the position of a Bragg angle because around that position efficiencies varied greatly; (ii) the reconstruction regions could be different and local variations occurred. Therefore the Bragg-angle reconstruction at different wavelengths from

different lasers is meaningless unless a laser with large tunability in wavelength, e.g., a dye laser, is available.

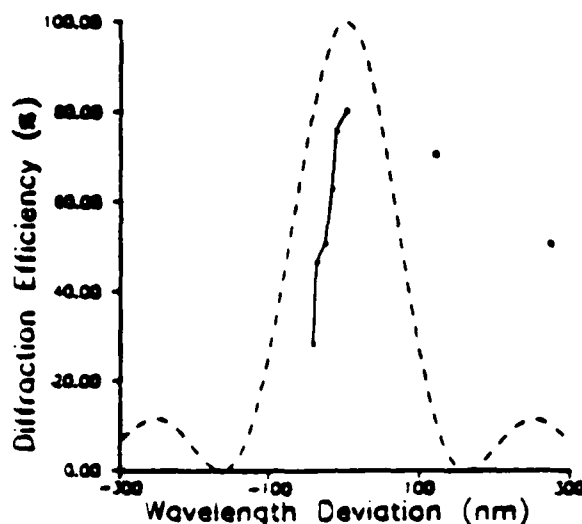


Figure 10.13 Spectral selectivity of a thick, phase grating of $E \approx 35 \text{ mJ/cm}^2$. The dashed line is for calculated values and the solid line for argon-ion laser lines at the Bragg angle; data shown in squares were taken for the maximum DE at a non-Bragg angle.

Measurements of the diffraction efficiency versus angle for both s- and p-polarization were also performed. At low exposures, diffraction efficiencies for both p- and s-polarization beams reconstruction show no appreciable differences as depicted in Fig. 10.11. At high exposures, however, diffraction efficiency of p-polarized light is about 3 to 5 times greater than that of s polarization (Fig. 10.14). Based on these observations, one might attempt to examine the possibility of using a HOE as a polarization beamsplitter. An appropriate choice of Bragg angle $\theta_{r,b}$ and product ΔnT can be shown to accentuate the diffraction efficiency of p-polarized light relative to s-polarized light. For example, Fig. 10.15 shows one possible solution in which the diffraction efficiency is zero for s-polarized light, and 100% for p-polarized light. This occurs when $\theta_{r,b} = 30^\circ$ and $\Delta nT = 0.44 \mu\text{m}$. It is also possible to form the opposite polarization accentuation as depicted in Fig. 10.16 with $\theta_{r,b} = 39.1^\circ$ and $\Delta nT = 0.705 \mu\text{m}$. It can also be shown that although solutions for

the above problems are not unique, the two examples given here are the only solutions feasible within the practical limits of ΔnT for DCG emulsion.

The grating design shown in Fig. 10.17 allows the diffraction efficiency of the hologram to be independent of the polarization of the reconstruction beam. The conditions for equal diffraction efficiency for both polarizations are $\Delta nT = 0.6 \mu\text{m}$ and $\theta_{r,b} = 36.7^\circ$. However, it is apparent that p-polarization in this case will have narrower angular selectivity than s-polarization. Comparatively, from the practical viewpoint, Fig. 10.11 provides a better way for polarization independent operation because angular selectivity is identical for both polarization and the maximum efficiencies are near 100% with only 2% deviation. All designs here are referred to $\lambda = 0.5145 \mu\text{m}$ and $T = 11.2 \mu\text{m}$. Extension to different wavelengths, various thicknesses of the recording medium, and a reflection-type phase hologram will be straightforward by employing the same design techniques based on Kogelnik's theory [10.35].

The rest of this section will propose a method to realize these designs and discuss possible applications of these devices.

To realize these designs one needs to control emulsion thickness and Δn . The spin-casting technique can generate uniform thickness over a large emulsion area. Speed of a spin is a determining factor in obtaining good uniformity for the success of this technique [10.41]. Special care is needed for the substrate preparation to reduce losses that are caused by defects on the substrate surface. The relationship between the concentration of ammonium dichromate and the incurred Δn must be examined to achieve the conditions for these designs.

Applications from the above design have commercial significance. For example, using a single holographic grating to replace a traditional bulky polarized beamsplitter is interesting to the field of optical data storage. The use of the polarization-independent grating as grating demultiplexers for fiber optics is another interesting application [10.42].

10.5 HOE Generating Multiple Equal Focusing Beams for Optical Interconnects

Many optical-interconnect applications require the optical elements to efficiently focus light. It is well known that CGH with the zone-plate-like structure can focus an incident beam. However, it is not easy for a single CGH to provide multiple equal focusing beams with high efficiency. In these cases, the patterns to be encoded are much more complicated than those of the binary phase gratings previously discussed and require expensive computer-aided design facilities to achieve the desired pattern. CGHs of this type also are often inefficient.

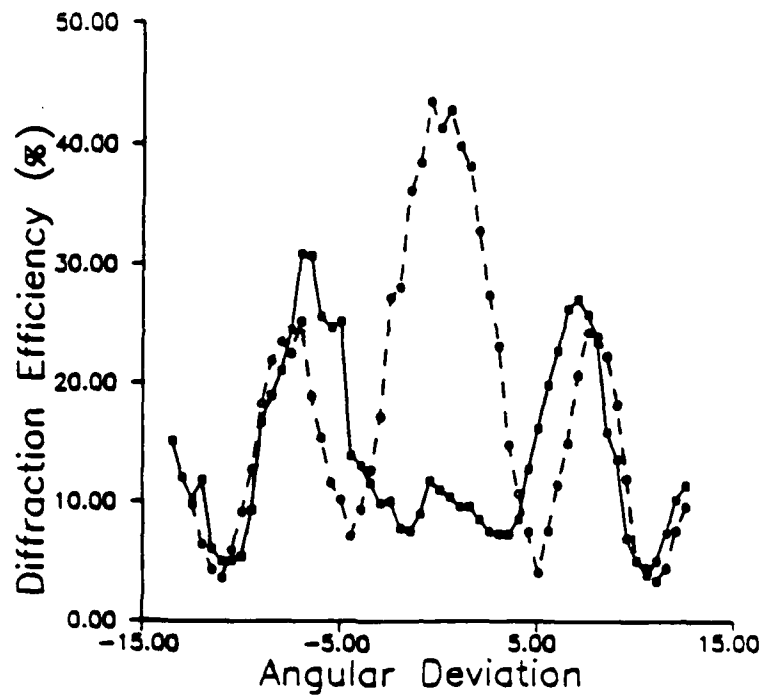


Figure 10.14 Angular selectivity of a thick phase grating of $E > 330 \text{ mJ/cm}^2$ shows high DE contrast around the Bragg angle; s polarization in squares and p polarization in circles.

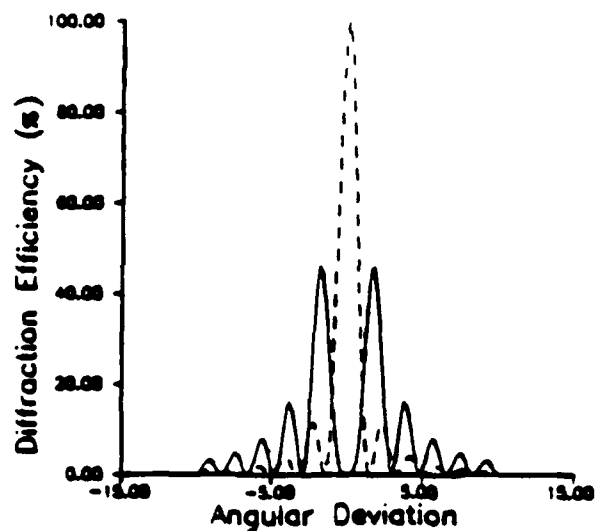


Figure 10.15 Calculated angular selectivity of a thick phase grating, showing DE of s polarization is zero and p polarization is one at the Bragg angle, where $\Delta nT = 0.445 \mu\text{m}$ and $\theta_{r,b} = 30^\circ$ are used for calculation.

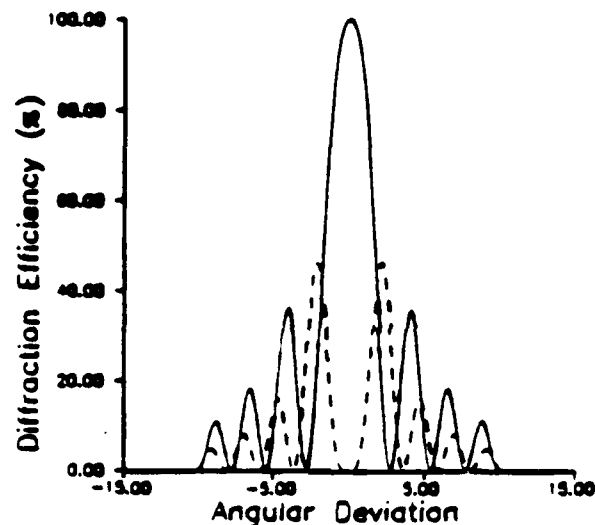


Figure 10.16 Calculated angular selectivity of a thick phase grating, showing DE of *s* polarization is one and *p* polarization is zero at the Bragg angle, where $\Delta nT = 0.705 \mu\text{m}$ and $\theta_{r,b} = 39.1^\circ$ are used for calculation.

In this section a simpler method is proposed to optically generate multiple-focusing beams and to efficiently record them. This method can bypass the involved and tedious direct CGH encoding but still can achieve the same goal of using a single element capable of generating an array of focused spots. Two steps are necessary to achieve this result: 1) generate multiple focusing beams that are equal in magnitude; and 2) efficiently record this object field.

Although the binary phase gratings previously described can be used to generate multiple equal beams, their diffraction efficiencies are limited by the accompanying higher-order diffraction beams that result from the thin grating characteristics. Referring to Fig. 10.4(b), side lobes may result that may have 15% to 40% of the total diffraction efficiency, depending on the number of beams used. The intensity distribution of these sidelobes can be simulated using the design work mentioned previously, although some tabulated work by Killat et al. [10.29], can provide a quick estimated value.

One way to increase the diffraction efficiency of the desired beams is to filter out the unwanted light by spatially filtering this pattern, depicted in Fig. 10.18, and

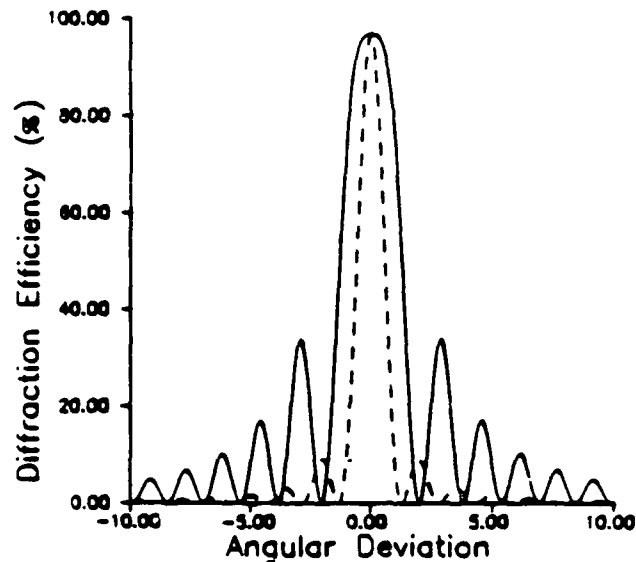


Figure 10.17 Calculated angular selectivity of a thick phase grating, showing DE of s and p polarization is the same at the Bragg angle, where $\Delta nT = 0.6 \mu\text{m}$ and $\theta_{r,b} = 36.7^\circ$ are used for calculation.

then holographically record the filtered wave onto the DCG. Therefore, the recorded element is expected to have focusing properties and high diffraction efficiency.

From the previous discussion on single-grating-thick holograms, we know that thick phase holograms have the advantage of high angular selectivity so that many holograms can be superimposed in the same space and read out separately if the total available index modulation is large. However, in general, to record multiple gratings some methods must be used to prevent cross coupling effects. For example, one way to avoid interaction between two gratings exposed simultaneously on a film is to angularly separate the gratings during exposure such that the Bragg angles are far from one another. Under these conditions each of the two gratings diffracts as if it were a single exposure grating for low exposures. Thus Eq. (10.25) still holds. However, at high exposure, the effect of saturation is such that to reach a given

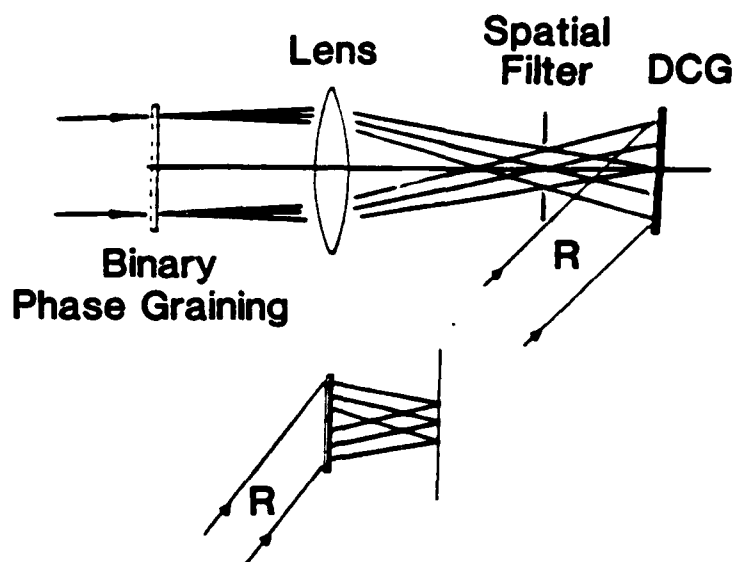


Figure 10.18 Optical layout of copying CGH onto DCG and its reconstruction.

index modulation in each of the two gratings, each grating must be exposed at a higher level than that necessary to make an equivalent single grating reach that same modulation level [10.40]. Of course, this compensation can only be carried out over a limited exposure range because the index modulation will be saturated as the number of gratings increases. In fact, each of the N gratings has a maximum modulation approximately equal to $1/N$ of the maximum modulation of a corresponding single grating [10.43].

Many theoretical or experimental efforts have occurred in the study of multiple *planar* grating effects, for example, work performed by Case [10.44], Kowarschik [10.45], and Kostuk [10.18]. However, few references describe the theory of simultaneously recorded, multiple-focusing gratings, not to mention an indicative beam ratio and exposure for optimum performance. Therefore, to obtain a HOE with maximum efficiency and signal-to-noise ratio, experiments were performed to optimize the beam ratio and recording exposure.

Figure 10.18 depicts the experimental layout. Diffracted fields from a nine-beam binary phase grating were focused by a lens of focal length 50 mm onto a DCG film through a spatial filter that filtered out unwanted higher diffraction

orders. The fields diverging from the spatial filter interfere with a reference beam to form an interference pattern at the DCG plate. The plate was placed so that the spot size of a reference beam is a little bit bigger than that of the diverging diffracted fields to ensure complete overlap. The surface normal of the film bisected the 36-degree angle formed between the reference and the diffracted fields. The working distance, the distance from a holographic plate to its focusing plane, was about 20 mm. Assuming diffraction-limited performance, the diameter of a spatial filter should be at least larger than $1.4mf\theta_d$, where m is the number of beams to be generated, θ_d is the diffracted angle of a grating, and f is the focal length of the lens.

Figure 10.19 shows the experimental results, indicating that diffraction efficiency is low for high beam ratios. Observations indicate that scattering noise and higher-order diffraction were greatly reduced as the beam ratio was increased.

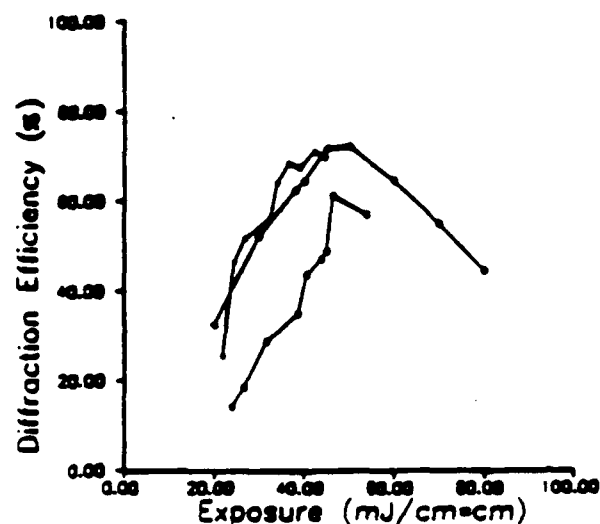


Figure 10.19 DE versus exposure for different beam ratios when copying multiple focusing beams onto DCG; data in circles are for $BR = 1.3$; squares for $BR = 10$; and the other for $BR = 2.2$.

Each individual reconstructed focused image was about $40\text{ }\mu\text{m}$ in diameter, with separations of $210\text{ }\mu\text{m}$ as measured by using the one-dimensional CCD Fairchild

Model 122 which has pixel resolution of $13\text{ }\mu\text{m}$. Intensity uniformity of these spots was good, as indicated by the demonstration of simultaneous switching of eight spots on a bistable ZnS interference filter. The maximum allowable variation of the switching power for this etalon array can be analyzed by referring to the work of Gigioli [10.46]. Nonuniformity occurs because the binary phase grating used to generate multiple beams before filtering was not well fabricated for $\lambda = 0.5145\text{ }\mu\text{m}$. However, the recorded HOE showed 70% diffraction efficiency, which was about 10% higher than the original binary phase grating; thus verifying the concept of improved efficiency with copying. Although one expects that higher diffraction efficiency than the one shown above can be achieved if a good CGH can be provided, the upper limit will be set by cross coupling effects. Figure 10.20a depicts the image of a nine-beam binary phase grating, and Figure 10.20b the magnified image of nine focused spots.

The major advantages of this copying technique are higher diffraction efficiency and compactness. A single copied holographic optical element can replace all other elements and reduce losses; moreover it improves efficiency, signal-to-noise ratio, alignment, and stability.

The previously mentioned switching experiment indicated that the HOE can be used to implement regular optical interconnects for optical computing such as the symbolic substitution algorithm. The same HOE also has potential use for interconnections in microelectronic systems. For example, one might couple a signal into a fiber bundle with the HOE and distribute these signals to several integrated circuits.

Summary

In this section we discussed several interconnect problems. A discussion of binary phase and volume gratings followed. A unique technique of implementing free-space optical interconnects was then described, which copied the diffracted field of a binary phase grating into DCG emulsion. The current experimental results show that the fabricated HOEs are capable of generating multiple focused beams with efficiency enhancement of at least 10% over the original binary phase grating. The HOEs can be used as an element for further investigating the feasibility of optical interconnects in VLSI technology.

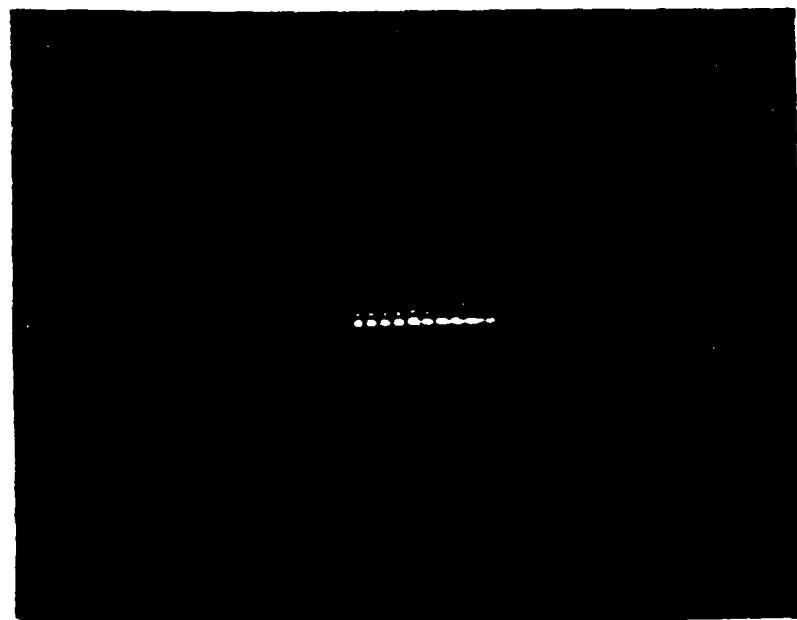
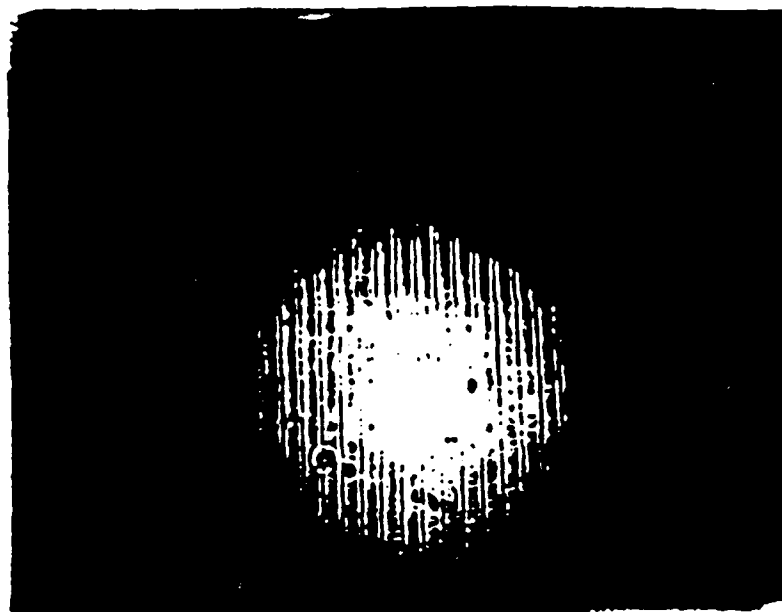


Figure 10.20 (a) A magnified image of a nine-beam binary phase grating; (b) A magnified image of nine focussed spots generated by a single HOE.

11. USE OF A HOLOGRAPHIC OPTICAL ELEMENT FOR FAN-OUT OF SIGNALS TO A GaAs ETALON

The distribution of signals is an important consideration when constructing an all-optical switching system. Holographic optical elements (HOEs) are promising for application in such systems because they are compact, rugged and inexpensive to produce. The light signals travel through free space, and the system takes advantage of the fact that many light signals can cross through the same space without affecting each other.

In this section the use of a HOE to produce multiple focused spots for switching a GaAs etalon is described. The HOE was made in the following manner. A computer-generated hologram was produced and used to generate a light field for a row of focused spots. This light field was spatially filtered to improve its characteristics and then recorded on a second hologram. The second hologram is the one used for the experiment described. The hologram was recorded in dichromated gelatin (DCG) using light with a wavelength of 514.5 nm. The HOE produces a 1×9 array of spots each about 40 μm in diameter and 190 μm apart [11.1].

The experimental setup is shown in Fig. 11.1. Modulated light at 514.5 nm is focused onto one side of a GaAs etalon using the HOE, and cw light at 871 nm is focused onto the other side using a 10 \times microscope objective. The output signal is the IR light reflected from the etalon. An etalon containing 299-Å multiple quantum wells was used in this experiment because its uniformity was the best of the etalons available at that time. We used two different wavelengths of light because the DCG used to make the hologram is not sensitive to infrared light. The best result achieved so far is modulation of the reflected IR beam by about 15% (Fig. 11.2). A transmitted beam should show much higher contrast.

The hologram used for this experiment has degraded from exposure to moisture in the air, and has an estimated efficiency of 15%. A new hologram of this type is 70% efficient, so the use of a new hologram should improve the fan-out efficiency by a factor of 4 or 5. Such holograms may be sealed in plastic to prevent any degradation.

The spot size of the IR beam is about 14 μm . Figure 11.3 compares the spot size of the IR beam and of the green-light beams. Table 11.1 shows the maximum power obtained so far for each focused spot.

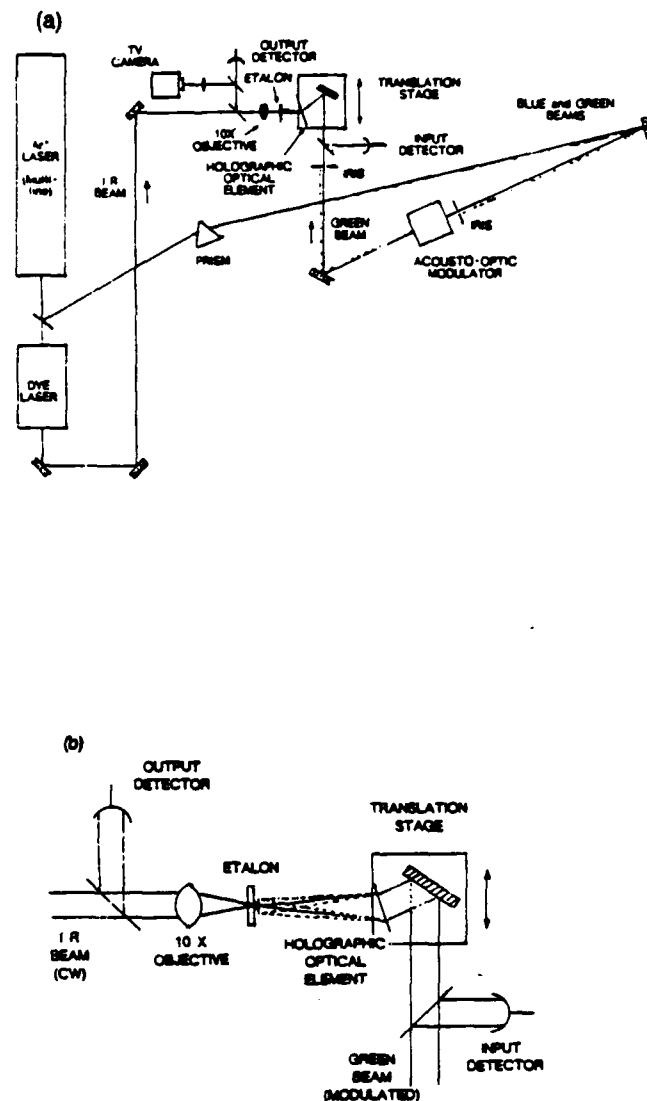


Figure 11.1 (a) Experimental setup for fan-out with the HOE. The 514.5-nm line of the Ar^+ laser is separated and modulated. The HOE and a mirror are mounted on a translation stage such that the focused spots from the HOE can be translated across the etalon without affecting the illumination of the HOE. A cw beam at 871 nm is reflected from the etalon and the output signal detected. (b) Detail of the optical system near the etalon.

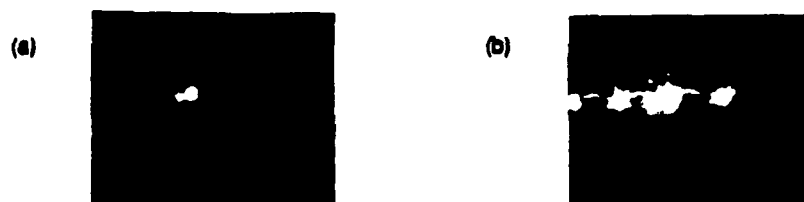


Figure 11.2 (a) Input signal (green beam, top) and output signal (bottom) with no IR beam, showing zero output. (b) Input signal (top) and output signal (bottom) with IR beam on, showing modulation of reflection of the IR beam by the green beam.

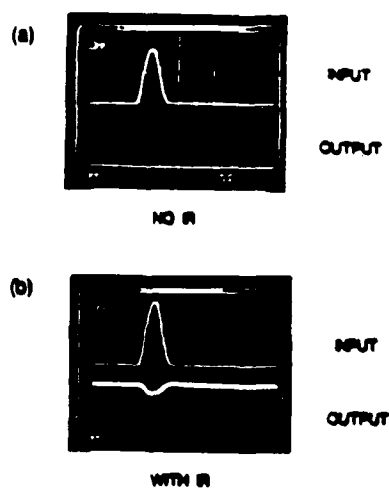


Figure 11.3 (a) Image of the focused IR spot reflected from an empty region of the etalon. Spot size is about $14\text{ }\mu\text{m}$. (b) Images of the green spots focused by the HOE and transmitted through an empty region of the etalon, with the IR beam superimposed. The spots from the HOE are about $40\text{ }\mu\text{m}$ in diameter.

Table 11.1 Maximum power obtained for each spot of the image from the HOE.*

Spot	1	2	3	4	5	6	7	8	9
Power (mW)	3.8	4.6	4.6	4.9	7.2	4.6	4.6	4.9	3.4

*A sealed hologram is expected to be 4 or 5 times more efficient.

B. COMPARISON OF DEVICES FOR OPTICAL SIGNAL PROCESSING

12. ACHIEVEMENTS VERSUS STATEMENT OF WORK: PROSPECTS FOR NONLINEAR DECISIONMAKING IN OPTICAL SIGNAL PROCESSING

12.1 Introduction

In this section we examine the characteristics of the various devices associated with our research and, with our original Statement Of Work as a base, describe our achievements and the problems encountered during the course of our work.

12.2 Interference Filters

12.2.1 Statement Of Work (4.1.1)

"Fabricate interference filters that have long-term stability and sufficient uniformity over a $5 \times 5 \text{ cm}^2$ area. The bistability switch-power shall not vary by more than 10%. The use of ZnS and ZnSe shall be required."

Achievements

Our work with interference filters, including design considerations and test results, is described in Section 2. We fabricated about 30 filters and thoroughly tested about 10 of these. Reasonably good uniformity was achieved, evidenced by the linear-transmission-peak wavelength as shown in Fig. 2.7, for five spots positioned in a 5-cm^2 square, as shown in Fig. 2.6.

A more concerted effort to achieve $<10\%$ variation in bistability switch-up power over a $5 \times 5 \text{ cm}^2$ was not made because:

- (a) Interference filters have a damage and long-term drift problem (Section 2.4). This problem apparently has been eliminated to a large extent by the Heriot-Watt group by growing the absorber layer *outside* the interference filter [12.1].
- (b) Optical bistability is highly sensitive to variations in thickness, reflectivity, laser power, detuning, etc., making it impractical for "infinitely" cascable computations (Appendix C).
- (c) While GaAs and ZnS devices have similar switch-up powers, the recovery time for GaAs is five orders of magnitude faster than that for ZnS. We were convinced we should concentrate on the GaAs etalon, which is our specialty, rather than work to optimize the ZnS interference filter.

12.2.2 Statement Of Work (4.1.1.1)

"Conduct a 25-beam demonstration unit. This unit shall demonstrate identical bistability anywhere on a 25-cm² unit."

Achievements

The simultaneous switching of 8 pixels with a separation $\approx 100 \mu\text{m}$ on a ZnS interference filter using a 1-D binary phase grating and a 50-nm-focal-length lens is described in Section 2.4. In Section 13.3 we report a pattern-recognition experiment utilizing a 2×9 input pattern to recognize three-spot patterns. This 18-beam pattern was shifted and recombined on one AND-gate filter and the output of that operation recombined with another shifted 18-beam pattern. Consequently, 36 logic gates were operated simultaneously, utilizing about $0.5 \times 4 \text{ cm}^2$ on each filter. This pattern recognition experiment exceeded the 25-beam target, and accomplished Statement Of Work 4.1.3 as well. However, the bistability loops were not identical, as discussed in Section 2.

12.3 GaAs Etalons

12.3.1 Statement of Work (4.1.2)

"Fabricate GaAs excitonically nonlinear etalon arrays."

Achievements

During this contract, our GaAs and MQW crystals all came from outside sources. Initially, our principal supplier was Art Gossard at AT&T Bell Labs. When he took a position at the University of California at Santa Barbara, the owners of the MBE machines blocked further collaboration between us. Other crystals were supplied by Keith Evans at Wright Patterson, Pallub Bhattacharya at the University of Michigan, and John Polky at Boeing. These crystal growers have gone to great lengths to supply our needs; we are very grateful to them. Nonetheless, we have at times waited months for samples; often the new samples have been designed to test new ideas rather than for use in device optimization or in demonstrations. This shortage has led us to obtain an MBE machine dedicated to GaAs/AlGaAs growth, with an emphasis on all-optical devices.

During this contract, we fabricated about 20 GaAs and GaAs/AlGaAs etalons, usually with external mirrors attached with a UV curing cement. Near the end of the contract, we obtained a sample grown to our design specifications with integrated mirrors (Section 4.2). An array of integrated Fabry-Perot devices was fabricated by photolithography and reactive ion etching (Fig. 4.10) with $5 \times 5 \mu\text{m}^2$ pixels separated

by 20 μm center-to-center. It is relatively easy to fabricate such arrays; in fact, we have etched structures with dimensions of only a few hundred angstroms in the course of our quantum-dot studies. Arrays of nonlinear decisionmaking elements can be defined by such pixellation or merely by forming an array of focused spots. Thermal considerations, recovery-time, and system considerations may dictate which method is better.

12.3.2 Statement Of Work (4.1.2.1)

"Use chemical and plasma etching techniques to achieve $\Delta L < 0.1\lambda/n_0$ over an area of one square centimeter."

Achievements

Section 4 describes the GaAs etalon development work. As reported on page 47, for the etalon with integrated mirrors, the resonance wavelength deviates less than three linewidths over the entire 1 cm^2 region. As its finesse was $\cong 18$, one can calculate the corresponding nonflatness as follows. The free spectral range corresponds to a change in length of $\lambda/2n_0$. Because the resonance linewidth is the FSR divided by the finesse, the corresponding change in length to the linewidth is $(\lambda/2n_0)/18$. Therefore, $\Delta L \leq 3(\lambda/2n_0)/18 = \lambda/12n_0$. So, our goal was achieved, not through etching techniques, but through the growing of integrated mirrors. This uniform sample worked very well in reflection mode but not in transmission mode, perhaps because GaAs was used in the mirror layers. We believe that the challenge to improve uniformity has not been addressed systematically because persons using MBE techniques are satisfied with the uniformity they have achieved for electronic and optoelectronic devices.

12.3.3 Statement Of Work (4.1.2.2)

"Construct GaAs pixel arrays with different dimensions by plasma etching. The characteristics of these different arrays shall be evaluated."

Achievements

See Section 12.3.1. Only a few etalons were pixellated because of a shortage of GaAs samples and the need to focus on problems discovered in device operation (Sections 12.3.4, 12.6 and Appendix C). Some system analysis was performed (Appendix C). Our reactive ion etching program has progressed steadily with the successful fabrication of nonlinear directional couplers [12.2] and quantum-confined cylinders, so pixellation can be performed whenever desired.

12.3.4 Statement Of Work (4.1.2.3)

"Investigate array density to determine limitations. The investigation shall include thermal limitations as well as other factors that limit packing density."

Achievements

The density of pixels is limited much more by the difficulty of optical interconnection than by the reactive ion etching process, which allows relatively easy creation of one-micron pixels separated by a few microns. Electron-beam lithography can be used to write 20-nm features. In fact, we have etched cylinders ≈ 60 nm in diameter separated by ≈ 400 nm. These arrays are too small for logic elements to couple into, but they illustrate that etching is not the limitation [12.3]. Thermal limitations are known to be critical (Section 12.6) with present devices, but the high-speed-device problem of cascading, unappreciated when the Statement Of Work was written, seemed more important to address.

12.3.5 Statement Of Work (4.1.2.4)

"Examine the problem of cascading several devices and seek solutions to this interconnectivity problem."

Achievements

Cascading of two GaAs etalons was achieved through optical bistability operation using microsecond pulses (Section 5). Cascading with nanosecond pulses failed (Section 6). Computer analysis revealed that gain vanishes as the pulse length is shortened to ten times the medium recovery time (Section 6). Alternatively, one can use two-wavelength NOR-gate operation with picosecond gain, but the output is shifted in wavelength, that is, it is unsuitable for use as an input to an identical device. A slight enhancement of the gain occurs through angle tuning of the switch beam (Section 7).

Cascading at GHz rates is a challenge. We concentrated most of our efforts on this problem rather than on the problem of achieving massive parallelism in optical interconnects. Some holographic optical element work was accomplished, including: optical copying of computer-generated holograms (Section 10); using the new hologram to simultaneously switch eight spots on a bistable interference filter (Section 10.5); achieving a 15% modulation of a test beam reflected from a GaAs etalon with a HOE producing a 1×9 array of spots $40 \mu\text{m}$ in diameter and $190 \mu\text{m}$ apart (Section 11). Ray Kostuk has constructed a state-of-the-art holographic facility; he is eager to address the interconnect problem when we reach the system phase. Related work, funded under other grants, was performed on photorefractive crystals [12.4, 12.5].

12.4 Application Demonstrations

12.4.1 Statement Of Work (4.1.3)

"Construct a demonstration array (3 pixel \times 3 pixel) to recognize basic patterns. Develop a demonstration optical system to recognize the occurrence of optical patterns. Beam division multiple-beam focusing and re-imaging shall be used in the development of an optical system."

Achievements

Accomplished; see 12.4.2.

12.4.2 Statement Of Work (4.1.3.1)

"Repeat the demonstration in paragraph 4.1.3 using a 4 pixel \times 4 pixel array. Conduct another demonstration by replacing the lens array with a holographic array."

Achievements

Section 12.2.2 and Section 13 describe a pattern-recognition experiment in which three-spot patterns were recognized using an array of 2×9 beams formed by two phase gratings (one 1×2 , the other 1×9) in tandem, and two interference filters, meeting the objectives of paragraphs 4.1.3 and 4.1.3.1.

12.4.3 Statement Of Work (4.1.3.2)

"Based on the outcome of the experiments in paragraphs 4.1.3 and 4.1.3.1, additional signal-processing goals shall be attempted."

Achievements

Faced with the problems of massive parallelism, namely those of obtaining sufficient laser power, beam dividing and combining, and filter non-uniformity, we chose two applications to signal processing where a few decisionmaking elements can have impact: 1) fingerprint identification by an associative memory using a nonlinear interference filter as a thresholding element in the correlation plane (Section 14); and 2) compare-and-exchange circuitry using four interference filters (Section 15).

12.5 Comparison of ZnS Interference Filters and GaAs Etalons

The higher speed of GaAs with little difference in bias power makes GaAs etalons more desirable than ZnS interference filters for our applications. However, the GaAs NOR gate is the only option with any promise for practical \geq GHz operation, and infinite cascability has not yet been achieved. A detailed comparison of the two devices follows in Table 12.1.

Table 12.1 Comparison of characteristics of ZnS IF and GaAs etalon devices.

	ZnS IF OB	OB	GaAs Etalon NOR-gate
Wavelength	514.5 nm	820-880 nm	820-880 nm
Switching Speed			
On	$\cong 1$ ms	40 ns	$\cong 1-3$ ps
Off	$\cong 1$ ms	40 ns	3-20 ns MQW 30 ps windowless GaAs
Switching Energy	10-100 μ J	0.6-60 pJ	0.6-60 pJ
Bias Power	$\cong 10$ mW	$\cong 10$ mW	- - -
Uniformity	Good but damage (solved at Heriot- Watt by external absorber)	< 0.1 λ/n_0 integrated mirrors	
Peak Transmission	50-60%	10-20%	10-20%
Contrast Ratio	4:1	10:1	10:1
Compatible with Diode Laser	Using other semiconductors	Yes	Yes
Maximum Bit Rate	$\cong 100$ Hz	$\cong 1$ MHz	Not yet cascable

12.6 Problems With Using Nonlinear Optical Devices

Nonlinear photonics can be divided into three areas: digital optical computing, optical neural networks, and photonic switching. In a search for the best applications for GaAs etalons we have considered each area and drawn some tentative conclusions.

What are some of the problems with using nonlinear optics for digital computing? Why has progress been so slow? The most thoroughly-analyzed device is the nonlinear etalon operating as a single-wavelength bistable element or as a logic gate with inputs and outputs at different wavelengths. Typical bistability holding power is 10 mW per device; typical energy per logic operation per device is 10 pJ. To utilize the massive parallelism of optics would then require, assuming 10^6 elements per plane, a formidable 10 kW/plane cw for bistability or 10 kW/plane average for a

1-GHz logic rate. If all of the power is absorbed in the devices, they must be spread out over 100 cm^2 to keep the heat load to 100 W/cm^2 . The problem of removing that much heat while accessing the etalon optically has not been addressed seriously. In part this is because there is still hope of reducing the energy per decision and in part because no acceptable solution to the high-speed cascading problem has been found.

How can the power and energy requirements be reduced? Enhancement of the nonlinearity is desired, but while inorganic semiconductors have been studied extensively, the search among organics has yielded no breakthrough so far. Quantum confinement increases the nonlinearity per unit volume but not the index change per absorbed photon. Perhaps a miracle material will be discovered, but there seem to be no fundamental-physics arguments to make us predict it. Another way to reduce the power and energy required is to reduce the size of the device. Microresonators of GaAs multiple quantum wells with nominal size of $1.5 \text{ }\mu\text{m}$ have been operated as NOR gates with switching energies as little as 0.6 pJ as a result of their reduced size and waveguiding. Such miniaturization leads to formidable problems in focusing onto large arrays of pixels and in re-imaging to the next array. A compromise may result in slightly larger devices demanding higher switching energies. Waveguides do not seem to be the solution either, because the energy required per device is no lower if the nonlinearity is resonant; the larger detuning from the band edge necessitated by the longer length results in a reduced index change, so reducing the power by increasing the interaction length is not possible. It is even more difficult to imagine massive parallelism with waveguides. Steady progress in the development of diode lasers and laser arrays is gradually improving the availability of laser power, but normally only one device can be powered by a single diode laser. We conclude that there is no reason to expect orders-of-magnitude reduction in power and energy requirements.

Another unsolved problem is the mode of operation. Optical-bistability and NOR-gate operations are the most often discussed, but both have serious problems. Optical bistability is a single-wavelength operation; the output and inputs can have the same wavelength, facilitating the cascading of devices, which is essential for digital optical computing. But bistability in GaAs etalons either is too slow or requires too much power, and is rather sensitive to fabrication and operation parameters (Section 12.7). Consequently, we abandon bistability for logic and reserve it for latching. A more promising mode for logic operation is NOR-gate, but achieving cascading remains a crucial problem for digital optical computing applications; the output has a wavelength different from that of the input.

What are some of the problems in achieving optical neural computing? The heart of practically every optical implementation of a neural network or real-time associative memory is a photorefractive crystal. Such crystals exhibit many attractive features, such as low-intensity (mW/cm^2) operation; same-frequency two-beam energy transfer; and self-pumped phase conjugation. Nonetheless, there are problems in applying them to complex systems, including difficulty in growing large homogeneous crystals; amplification of scattering (fanning); a limit on the number of images that can be stored in a single crystal; response time constraints; instabilities; angular dependence of energy transfer; and freezing in the "learned" gratings. Many system designs also include nonlinear etalons for decisionmaking or for controlling sigmoidal transmission to achieve backward error propagation. There are no photorefractive crystals that work well at the operating wavelengths of GaAs etalons. Because both GaAs etalons and ZnSe interference filters require kW/cm^2 they are incompatible with efficient photorefractive crystals which require only mW/cm^2 , unless one is performing whole-image decisionmaking. Nonthermal, slow nonlinear devices, based perhaps on nipi structures, would be much more compatible with photorefractive crystals. Unless slow, low-power, all-optical devices can be developed, electro-optics devices such as the SEED (self electro-optic effect device) may be the only practical solution. Such devices may be able to preserve the phase, as is often required. Otherwise, optics will perform only the interconnect function; this may be crucial (Section 10). The electronic competition is also formidable in the neural-net area, and rapidly increasing.

Finally, nonlinear optics can be applied to photonic switching for demultiplexing, time re-ordering, high-speed routing, sampling, etc. Optics leads electronics by far in short-pulse generation, transmission, and diagnostics. Several all-optical switches that open or close in a picosecond have been demonstrated, with recovery times varying from 1 ps to several nanoseconds. Multiple quantum wells and optical glass fibers have been fabricated into picosecond switches, and many organic materials exhibit the appropriate response times. Switching seems to be the area of nonlinear photonics where nonlinear optics makes unique contributions, accomplishing feats electronics is incapable of matching. Clearly, the disadvantages of high switching energy and two-wavelength operation are not nearly as serious for switching as for optical computing.

The division of nonlinear photonics into neural computing, digital computing, and high-speed switching emphasizes their different time regimes: milliseconds, nanoseconds, and picoseconds, respectively. Progress in decisionmaking devices for one regime may not, therefore, carry over to the other two. Major advances in photonic switching, however, may revolutionize some aspects of digital computing, for

example, by making possible an ultrafast optical preprocessor or by removing a bottleneck in an electronic computer.

C. OPTICAL SIGNAL PROCESSING USING ZnS INTERFERENCE FILTERS FOR NONLINEAR DECISIONMAKING

13. Pattern Recognition and Symbolic Substitution [13.1]

13.1 Introduction

Interference filters with ZnS or ZnSe spacer layers exhibit thermal refractive optical bistability. These IFs show potential for application as optical switches in digital optical computing. In spite of their slow recovery time (on the order of a millisecond) compared to that of GaAs etalons (on the order of a nanosecond), interference filters are still useful for implementation of algorithms and for demonstrations of optical computing schemes, especially at the early development stage. As described in Section 2, these filters operate in the visible wavelength range at room temperature, and are relatively easy to fabricate with reasonable uniformity, good stability, contrast, and fan-out.

In symbolic substitution logic (SSL), information to be processed is encoded in patterns [13.2]. The two main stages of SSL are pattern recognition and symbol scription. Simultaneous parallel computations are performed through recognition of a specific pattern and then substituting another pattern for the one recognized. An example is shown in Fig. 13.1(a), in which a binary "zero" is a bright spot below and dark spot above and vice versa for a "one". This dual rail "one" and "zero" logic format encoding simplifies thermal stabilization, because the heat incident on the device is always the same. Copies of this input are made; one then can make relative shifts of the copies to identify the occurrence of one or more particular patterns. The identification decisions are made by an array of nonlinear optical logic gates on which the input and shifted copies are superimposed. The output answers are used to generate a new pattern according to predefined substitution rules, thus completing the substitution stage.

In this chapter we present the design of a shifter, which constitutes an important element for performing SSL. This presentation is followed by a description of the IFs to be used in the experiments. Then, three optical computing experiments using ZnS or ZnSe interference filters as logic gates are described: 1) a simple pattern recognition; 2) a three-spot recognition; and 3) a complete SSL demonstration. Because SSL can be used to realize any Boolean logic operator, the principle of an optical one-bit full adder is described and demonstrated. Finally, some practical problems are addressed.

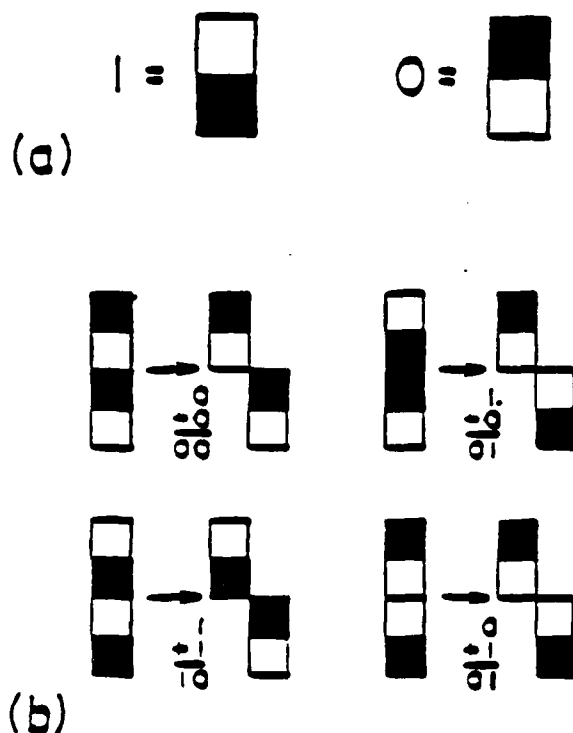
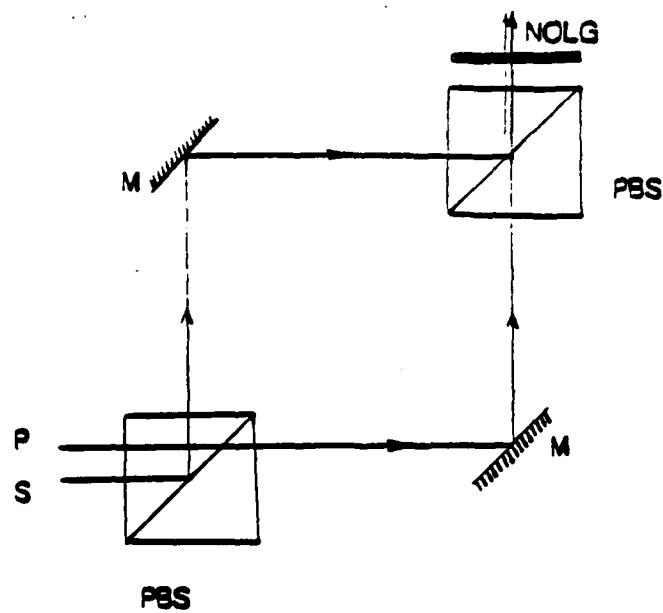


Figure 13.1 (a) Dual-rail binary encoding of logic "ONE" and logic "ZERO." (b) Rules of addition by CSL.

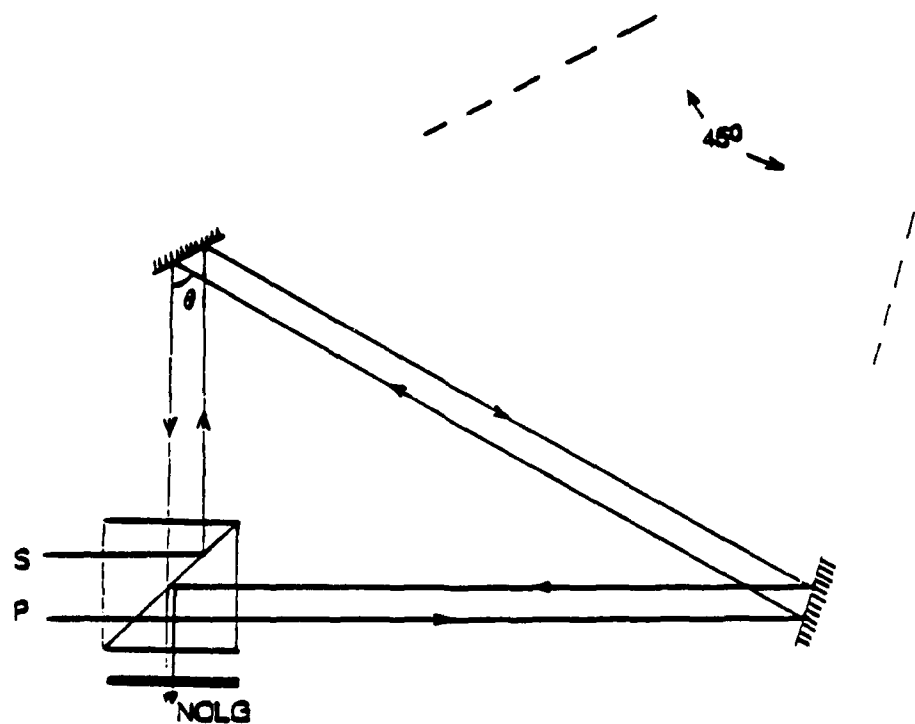
13.2 Component Design and IF Performance

The shifting operation is essential for both pattern recognition and substitution. From the experiments described later come ideas for the design of a shifter aimed at improving system accuracy and optimizing compactness.

The Mach-Zehnder-interferometer-type shifter used in our experiments results in unequal optical pathlengths for both inputs to the AND-gate array, as shown in Fig. 13.2(a). This property becomes a critical problem when one uses very fast gates driven by fast pulses, such as picosecond pulse trains incident on an array of GaAs etalons. One possible solution is to insert a compensation plate in one arm of the interferometer to make both paths equal. An alternative is to use a Sagnac interferometer, as depicted in Fig. 13.2(b), in which both arms have equal paths.



(a)



(b)

Figure 13.2 (a) Shifter of a Mach-Zehnder-interferometer type showing unequal paths. (b) Shifter of a Sagnac-interferometer type showing equal paths. PBS: polarization beam splitter; NOLG: nonlinear optical logic gate.

An innovative shifter design is shown in Fig. 13.3(a); it is a modification of the Sagnac interferometer. This shifter employs two half-pentaprisms, a plano-convex lens and a mechanical translator. It can be used to shift and to focus beams onto bistable logic gates. In addition to its compactness, this shifter offers two major advantages. First, the relative shifting of an input pattern can be done just by translating one prism parallel to the contacting side along line BD, as shown in Fig. 13.3(a). The second advantage is that the shifter can be used as a lateral shearing interferometer to test the collimation quality of an incident wavefront, which helps alignment greatly.

A 45° -polarization beam-splitting coating is deposited on the movable prism so that s-polarization beams will reflect and p-polarization beams will transmit when they are incident on the coating at 45° . Both s- and p- polarization beams, however, will transmit at normal incidence after bouncing back from mirrors. A suitable index-matching fluid is used between the prisms to reduce surface reflection losses. The choice of fluid is important for matching the effective refractive index of thin-film stacks and the refractive index of the prism glass. The effect of the small index-mismatch on polarization-splitting is expected to be mild. (A thorough study of this effect will be integral to the coating design. Such a study is beyond the scope of our discussion, and is left for future work.) A plano-convex lens is assembled with the fixed prism to provide focusing spots on the plane of bistable logic gates. Fabrication of the shifter can be summarized as follows. Refer to Fig. 13.3(b).

1. Cut a pentaprism into two symmetric halves with the following dimensions: $AO = u$, $AC \cong 2.165u$, $BD \cong 3.414u$, $CD \cong 0.586u$, $BE \cong 2.828u$ where u is an arbitrary measurement unit.
2. Deposit a 45° -polarization coating along BE.
3. Cement a plano-convex lens to plane AB. The lens has an aperture diameter equal to $2u$ and focal length $f > 6.828un_p + t_l n_l$, where t_l is the thickness of the lens and n_p and n_l are the refractive indices of glass materials for the prism and the lens, respectively.
4. Put a suitable index-matching liquid along plane BD.

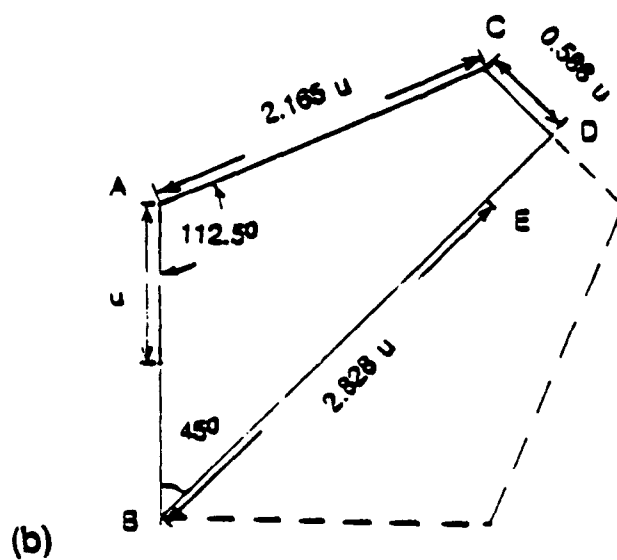
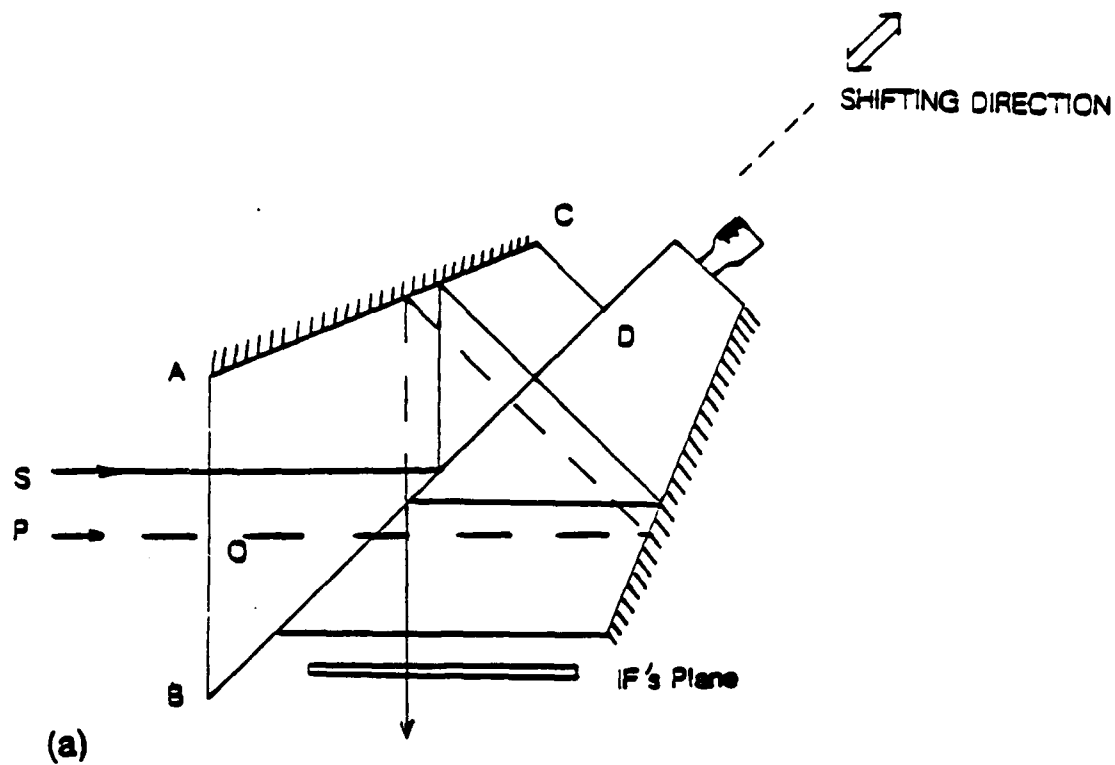


Figure 13.3 (a) Compact shifter of a modified Sagnac-interferometer type showing equal paths. (b) Detailed layout of pentaprisms of the new shifter.

Shown in Fig. 13.4(a) is a basic subsystem used to perform optical computing such as SSL, which incorporates a source, a plano-convex lens L and the above shifter. A source may consist of a laser diode phased array or an ion gas laser with a binary phase grating. The plano-convex lens is oriented such that the planar side faces the source, to reduce aberrations. The explicit layout and its y - \bar{y} diagram are shown in Fig. 13.4(b). In order to generate well-resolved focused spots, the resultant aberration of this optical system should be less than $\lambda/4$, according to Rayleigh criterion. The design of a shifter based on commercially available optics is shown, and aberrations are calculated up to third-order. As seen in Fig. 13.4(b), the shifter is a thin plano-convex lens with thickness t_l combined with a glass block with thickness t_p . The latter can be regarded negligible as t_p is much greater than t_l . Assuming the incident beam is collimated, the third-order aberrations of the shifter itself are calculated and summarized as follows.

$$W_{040} = \frac{1}{8} y_p^4 \phi^3 \left[A - B + C + D - \frac{(n+1)(n-1)^2}{n^3} c t_p \right] \quad (\text{spherical}) \quad (13.1)$$

$$W_{131} = \frac{1}{2} y_p^2 \phi^2 M \left[E - F - \frac{(n^2+1)(n-1)}{n^3} c t_p \right] \quad (\text{coma}) \quad (13.2)$$

$$W_{222} = \frac{1}{2} \phi M^2 \left[1 - \frac{(n+1)(n-1)^2}{n^3} c t_p \right] \quad (\text{astigmatism}) \quad (13.3)$$

$$W_{220} = \frac{1}{4} M^2 \frac{n-1}{n} c t_p \quad (\text{field curvature}) \quad (13.4)$$

and

$$W_{311} = \frac{1}{2} \frac{-M^3}{y_p^2} \frac{(n+1)(n-1)}{n^3} c t_p \quad (\text{distortion}), \quad (13.5)$$

where y_p is the marginal ray height at the first surface of the shifter; optical power $\phi = c(n-1)$, with c the curvature of the first surface; M is the Lagrange invariant; n is the refractive index of the prism material; and A, B, C, D, E, F are defined as:

$$A \equiv \frac{n+2}{4n(n-1)^2}, \quad B \equiv \frac{n+1}{n(n-1)}, \quad C \equiv \frac{3}{4} + \frac{1}{2n}, \quad D \equiv \frac{n^2}{4(n-1)^2}, \quad E \equiv B/2, \quad F \equiv 1 + \frac{1}{2n}.$$

Conditions for zero spherical aberration, zero coma and zero astigmatism are expressed as

$$W_{040} = 0 \quad \text{when } ct_p = \frac{n^3(n^3-2n^2+2)}{n(n-1)^4(n+1)}, \quad (13.6)$$

$$W_{131} = 0 \quad \text{when } ct_p = \frac{n^2(-n^2+n+1)}{(n-1)^2(n^2-1)}, \quad (13.7)$$

and

$$W_{222} = 0 \quad \text{when } ct_p = \frac{n^3}{((n-1)^2(n+1))}. \quad (13.8)$$

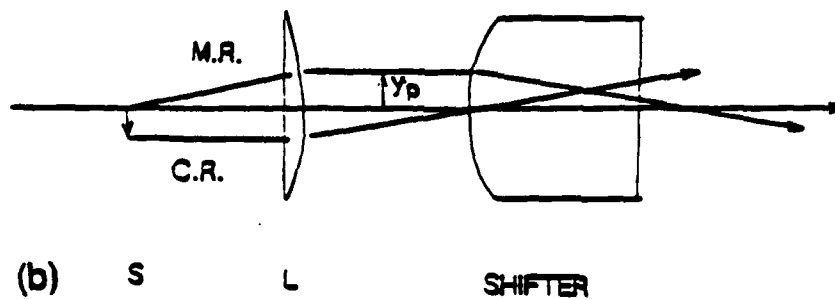
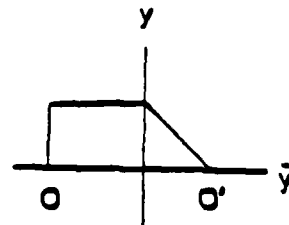
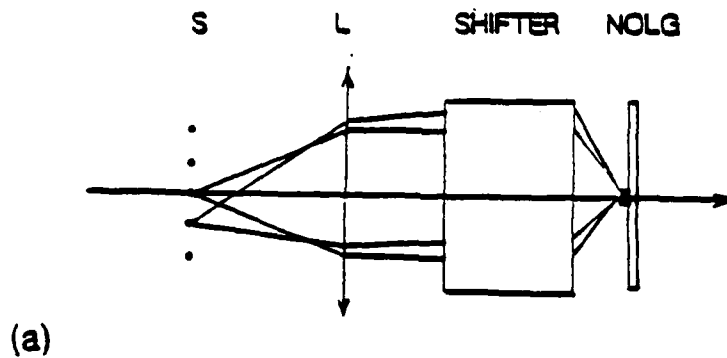


Figure 13.4 (a) Optical system for performing shifting and focusing required by symbolic substitution logic (SSL). (b) Optical layout and its y - \bar{y} diagram of the system in (a).

Three curves of n versus ct_p are plotted in Fig. 13.5, each representing $W_{040} = 0$, $W_{131} = 0$ and $W_{222} = 0$. Apparently there is no condition such that these aberrations vanish simultaneously. However, there are points where two of three curves intersect. That is, either $W_{040} = W_{222} = 0$, or $W_{131} = W_{222} = 0$ simultaneously at some particular sets (n, ct_p) as follows:

$$(n = 2, ct_p = 2.67) \text{ for } W_{040} = W_{222} = 0.$$

or

$$(n = 1.37, ct_p = 8) \text{ for } W_{131} = W_{222} = 0.$$

Unfortunately, glass materials exhibiting these values exactly are not available. One possible alternative is to use glass SF-59 ($n \cong 1.955$), which brings the spherical aberration to zero with some residual astigmatism and coma. Based on our preliminary analysis, future design work that takes advantage of the commercial lens design program, for example, Code V, can optimize system performance.

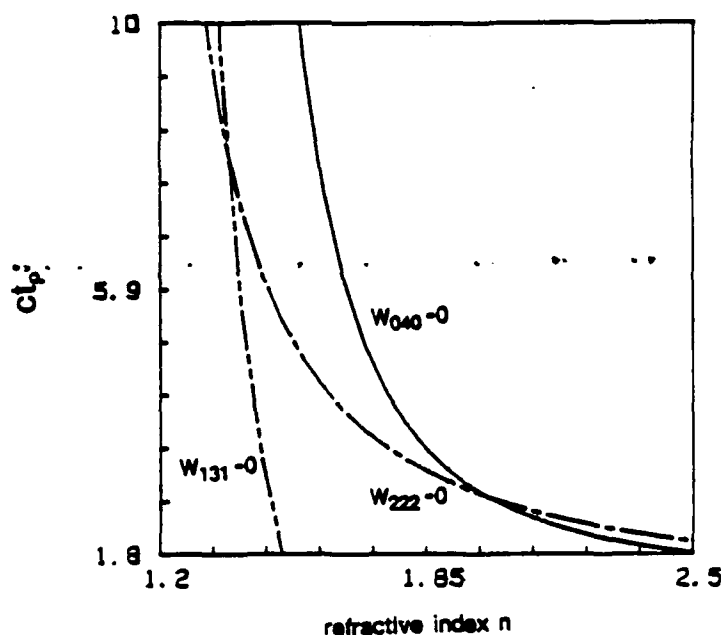


Figure 13.5 Conditions for zero spherical aberration, coma and astigmatism.

The interference filters used in our experiments, labeled IF1, IF2, IF3 and IF4, have the characteristics shown in Table 13.1.

Table 13.1*

	IF1	IF2	IF3	IF4
formula	$G(HL)^4 6H(LH)^4 G$	$G(HL)^4 4H(LH)^4 G$	$G(HL)^4 8H(LH)^4 G$	$G(HL)^5 6H(LH)^5 G$
$T_{\text{peak,linear}}$	52%	54%	54%	40%
λ_{peak}	511 nm	514 nm	514 nm	514nm
FWHM	19 Å	17 Å	17 Å	13 Å
P_T (min.)	25 mW	11 mW	9 mW	16 mW
$T_{\text{nonlinear,ON}}$	34%	31%	25%	12%
substrate	Kodak glass	SF-6	Kodak glass	Kodak glass

*A fan-out of 4 and a contrast of 2-2.5 are obtained when the ZnS IF2 is operated with a holding beam $\cong 45$ mW, and IF3 has fanout $\cong 1.4$ and contrast $\cong 3.5$.

13.3 Pattern Recognition

13.3.1 Simple Pattern Recognition

Pattern recognition, the first step in symbolic substitution logic, has been executed using ZnSe interference-filter type nonlinear bistable devices with linear optical components. The setup for this process is shown in Fig. 13.6, where the input is encoded on a mask of 35 mm film with transparent and opaque spots. The goal is to identify the pattern formed by two bright spots located next to each other. This identification is accomplished as above by shifting horizontally one pixel; the two shifted patterns are superimposed with a beam combiner. Each spot in the combined pattern can have one of three light levels: no light, one unit of light, or two units of light. The logic decision is made by focusing the combined pattern onto a 2×3 AND-gate array using a Nippon Sheet Glass fly-eye lens array. This results in 1 mm pixel spacing on a ZnSe interference filter whose structure is $G(HL)^4 6H(LH)^4 G$. The switching power per pixel is about 20 mW. The output is a bright spot in the right-hand pixel wherever the input has bright spots adjacent horizontally. Note that all the locations of this pattern are identified by a single logic operation.

independent of the array size; this is a clear utilization of the parallel nature of optics.

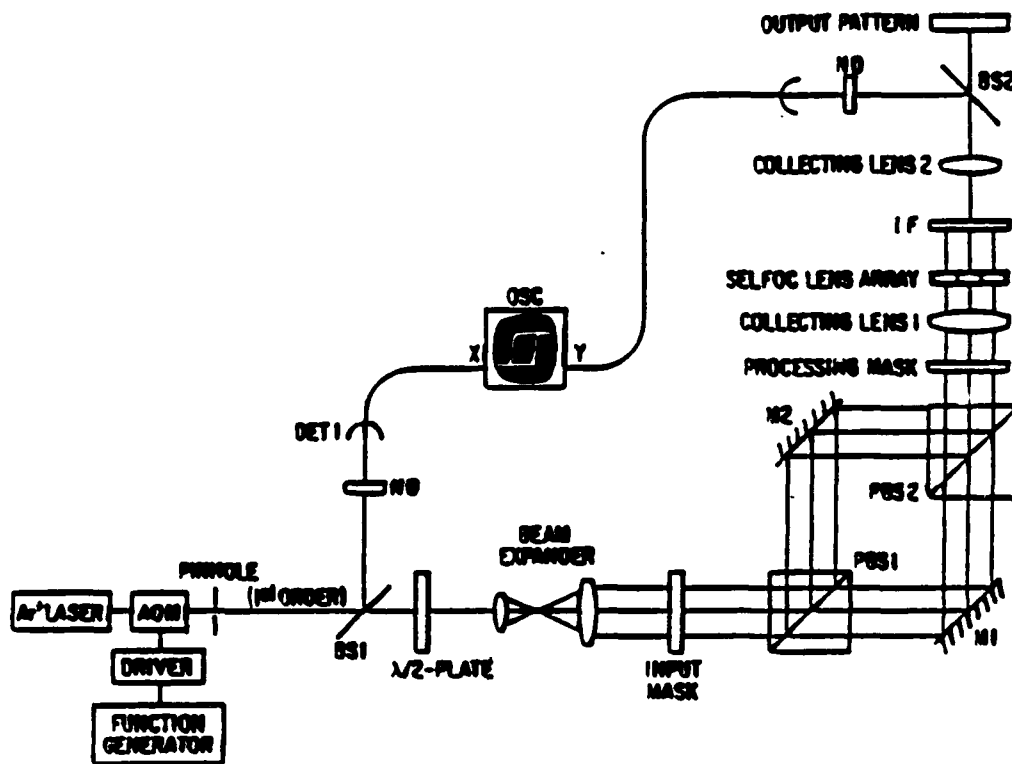


Figure 13.6 Experimental arrangement for the identification of two bright spots located next to each other horizontally (see text).

Referring again to Fig. 13.3 and Fig. 13.4 for SSL, the orthogonal polarization of the two input beams prevents interference between them. After receipt of the input, the time required to complete the pattern recognition is the switching time for the gates and the propagation delay from the input mask to the gate array. Thus, a large array of fast optical gates should give a large throughput.

13.3.2 Three-Spot Pattern Recognition

The goal of the three-spot pattern recognition experiment was to identify the locations of simple patterns consisting of three beams, for example the V and Γ shown in Fig. 13.7, in an input array. In this case a 2×9 beam array was used, with several randomly-selected beams blocked. Such a three-spot pattern could be

recognized using a single IF as a three-input AND gate, but here two successive two-input AND-gate operations are performed to emphasize the capability of cascading the nonlinear IFs.

One-bit addition by the technique of symbolic substitution [13.2, 13.3] has been demonstrated previously using IFs, showing that very simple patterns can be useful in digital optical computing [13.4]. The present experiment shows more parallelism and the potential for nonlinear decisionmaking in optical pattern recognition.

13.3.3 Experimental Apparatus

The apparatus is diagrammed in Fig. 13.7. A Coherent Innova-20 cw argon laser operating at 514 nm was the light source. The beam was modulated by a rotating half-wave plate and polarization beam-splitter, with a maximum transmission of $\approx 76\%$ and minimum of 3%. Two binary phase gratings, one that produces two and the other nine approximately equal beams, are oriented orthogonally and used in tandem to divide the 2.3-watt input beam into 2×9 beams. Details about these phase gratings are discussed in Section 10. Patterns are then encoded into the beam array by blocking selected beams. The power in each beam is 58 mW within 10%, after reducing the two strongest center beams by about 30%.

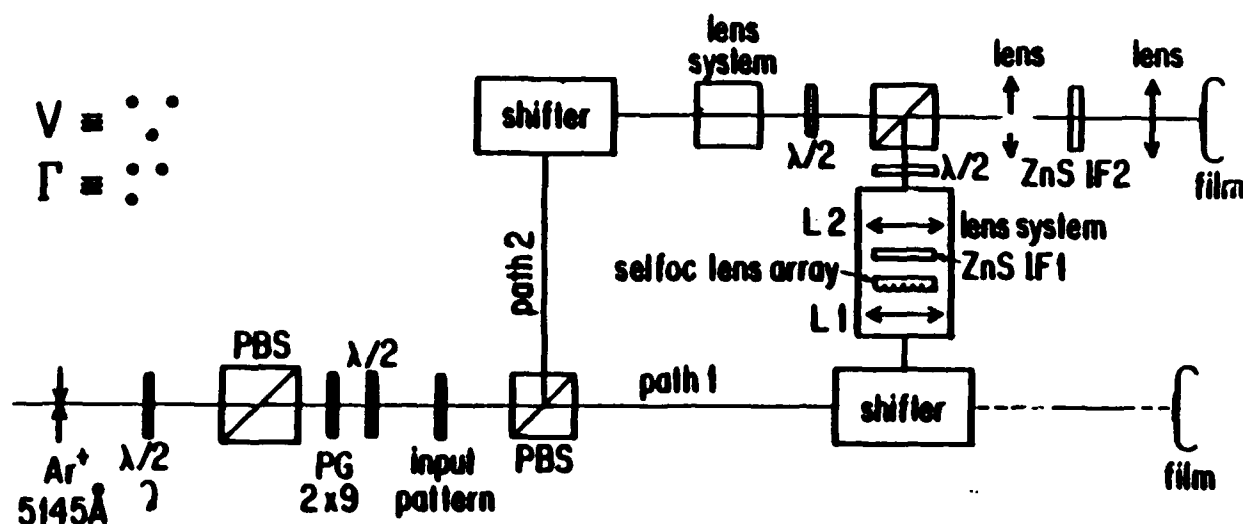


Figure 13.7 Experimental arrangement for digital pattern recognition.

ZnS interference filters IF3 and IF4, used as nonlinear decision-making devices, have characteristics summarized in Table 13.1. With appropriate scaling of the input intensities or holding beam intensities they require two or more inputs to switch to high transmission; therefore, these nonlinear devices can be used as bistable AND gates. In multiple-beam experiments the focusing is often imperfect, so typical switch-up powers are 20 to 40 mW on the filter and the nonlinear transmission is typically 10% to 25%. However, the contrast is ≈ 3.5 , permitting fan-out and cascading. Other system losses, mainly attributed to the optics for beam modulation and to phase gratings for multiple beam generation, limit the number of AND-gate operations per watt to 5 or 10, thereby limiting the parallelism currently achievable. The filters were grown for near-perpendicular operation at 514.5 nm, minimizing the complications of parallel operation on a tilted device.

Patterns were tested by sequentially identifying portions of the pattern. This operation was performed in conjunction with the shifters. The array of beams was first split into two arrays. One of the arrays was then shifted, and recombined with the unshifted array upon the IF1. In this way a portion of the pattern was tested. For example, a diagonal shift was required to identify the left portion of a V. Since the modulation is slow the path difference between two arms is negligible; therefore either a Mach-Zehnder-interferometer or a Sagnac-interferometer setup could be used as a shifter.

Each lens system in Fig. 13.7 focuses each of the input beams so that the parallel logic operations can be performed. The spots on the filters were about $75\text{ }\mu\text{m}$ in diameter and 1 mm apart. On IF1 the shifted input pattern and original pattern were superimposed by the shifter and focused by the lens system. On IF2 the pattern, appropriately shifted to search for the third spot, was focused on the re-imaged output of IF1. The lens systems consisted of two lenses, L1, L2, and a Selfoc lens array, each with focal length of 254 mm, 100 mm and 5.5 mm respectively. The combination of L1 and the Selfoc lens array was used to provide enough intensity for AND-gate operation on IF1. An appropriate mask was put between the Selfoc lens array and L1 to define the useful overlapped region. Lens L2 acted as a relay lens to pass signals to the next stage IF2. To minimize aberrations in the system, the lens system in path 2 was symmetrical to that of path 1, except that in our simple demonstration no logic operation was needed in this path for recognizing the third spot.

13.3.4 Experimental Results

The first step in recognizing a V is shown in Fig. 13.8. The upper left region shows the input pattern, encoded in a 2×4 format. The unshifted pattern and the pattern shifted diagonally down to the left are superimposed in the upper right region. The output of the AND-gate operation, shown below the input, gives a bright spot at the bottom of the V. Of course, this only shows that the right part of the V is present. A second AND operation, with the pattern shifted diagonally down to the right, completes the V recognition.

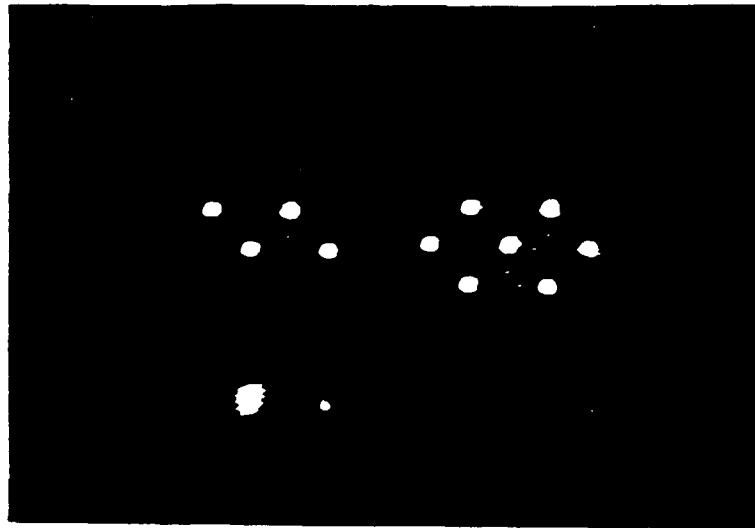


Figure 13.8 Recognition of "V." Input pattern is in the upper left region. Shifted pattern is in the upper right region. The result is in the lower left region, showing the occurrence of the right side of a V in only one location.

Recognition of three Γ 's in a 2×9 input pattern is shown in Fig. 13.9. Clearly, any three-spot pattern or simultaneous occurrences could be used to trigger an output response from this all-optical coincidence circuit, even though the experiment is slow and power hungry.

13.3.5 Complete SSL Demonstration

A simple but complete symbolic substitution has been achieved as shown in Fig. 13.10. The desired pattern is the simultaneous occurrence of bright spots in the lower left-hand and upper right-hand corners of an arbitrary 2×2 array. When this pattern occurs, the symbol-scription SSL stage generates an output pattern consisting of a bright top row and a dark bottom row. This step is accomplished by

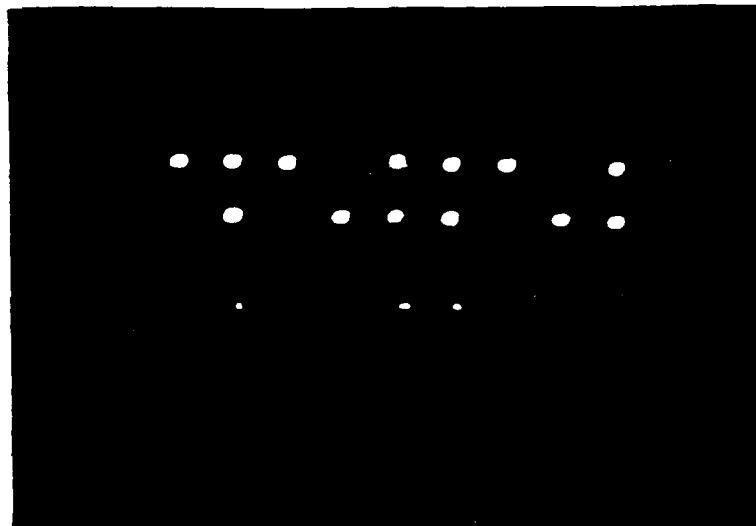


Figure 13.9 Recognition of "Γ." Input pattern is the upper two rows. The output of the two successive AND-gate operations used to look for Γs is shown below, clearly indicating the presence of three Γs.

an AND-gate operation using the output of the recognition stage with the strong holding beams. If the desired pattern is not present, the output is low, the result of lower branch transmission.

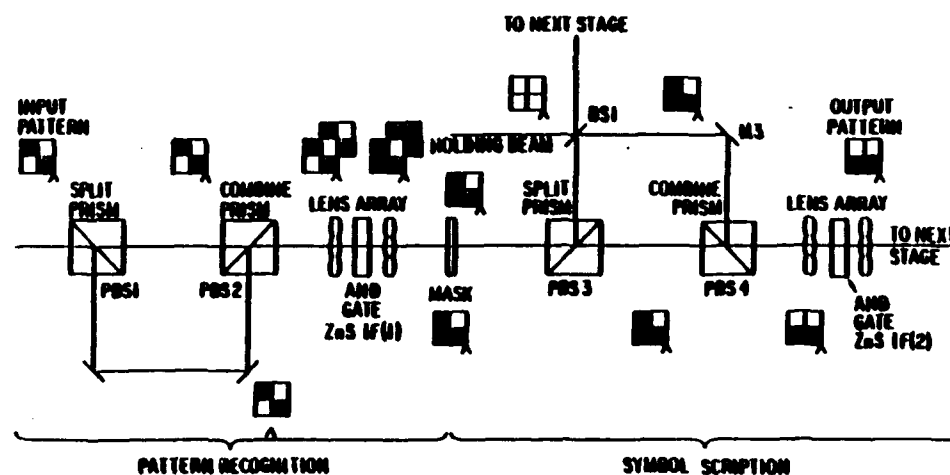


Figure 13.10 Simple but complete SSL experiment using two ZnS-IF AND gates. The marker "x" indicates the position of the bottom right-hand corner of the 2×2 array if no shifts are performed.

The desired diagonal pattern is recognized by dividing the input into two equal parts and recombining them on a ZnS AND gate after shifting one pattern up one square and to the right one square. If both squares on the desired diagonal are bright, there will be two units of light, enough to turn "on" the AND gate in the upper right square of the 2×2 input array. A mask is used to avoid the ambiguity that occurs if this 2×2 array is embedded in a larger array and the same pattern recognition executed. Although the computations accomplished by these experiments are simple, they illustrate pattern recognition, cascading, and symbol scription. In addition, the experiments require considerable expansion of single-beam techniques, including beam division, multiple-beam focusing and re-imaging, nonlinear etalon uniformity and stability. Techniques that are easily extendable to massive parallelism (100×100 arrays or greater) were used wherever possible (though this is not yet the case for the symbol-scription stage reflected in Fig. 13.10).

13.4 One-Bit Adder Experiment

An element essential to a digital optical computer designed for numerical computations is an all-optical full adder. A one-bit full adder is examined here.

Addition can be performed by SSL. The substitution rules shown in Fig. 13.1(b) are for the addition of two one-bit numbers: zero plus zero is zero with no carry, one plus one is zero with carry of one, etc. These rules are constructed to be operated in parallel. However, the logic operations required to implement the rules are not unique. One possible implementation scheme is shown in Fig. 13.11, in which the four input pixels (2 bits) are labeled 1, 2, 3, and 4 from top to bottom. The top part of the carry is bright only when both 1 and 3 are bright, so a simple AND gate is needed. The lower part of the carry can be determined by an OR operation. The AND and OR operations can be performed on the same filter if the holding beam intensities are appropriately different. Thus, one-bit addition can be accomplished with three channels: channel 1 is for both parts of the carry bit, channel 2 is for the upper pixel of the sum bit, and channel 3 is for the lower pixel of the sum bit. An alternative for channel 1 is to add an additional "zero" before performing an AND gate operation, instead of varying holding beam intensities for AND and OR gate operations on the same IF. Both implementations for channel 1 are equivalent, except that the latter requires only an AND-gate operation. An example, in which binary addition of one plus one is explicitly executed, is shown in Fig. 13.12. Each channel should have the same optical path, to avoid delay error when the outputs are combined. Note that there is only one bright spot in the sum and one in the carry; therefore, one spot is the complement of the other. Thus a

good inverter would simplify the nonlinear decision making greatly. Extension to an n -bit adder is simple, with the same implementation except that $n \times 4$ pixels and $n + 1$ iterations are required to achieve the final answer. A block diagram of an n -bit adder using SSL is shown in Fig. 13.13. The carries will automatically propagate for symbolic-substitution n -bit addition; thus, an n -bit full-optical adder is feasible.

13.4.1 Experimental Realization

To implement a one-bit adder, 1×4 collimated equal beams are required if the input bits are defined by blocking the beams appropriately. Implementation has been accomplished by first shifting parallel one mirror of a Sagnac interferometer to obtain the desired separation for 1×2 equal beams and then repeating this operation to produce 1×4 equal beams. The Sagnac interferometer ensures equal optical paths. Half-wave plates can be used to balance the powers of the 1×4 beams. According to the scheme in Fig. 13.11, one must copy the pattern, shift one copy up two squares, and recombine them on ZnS interference filters which act as AND and OR gates for different channels. The modulated holding beam is held so that two inputs are required to form logic "one" for an AND gate while only one input is required for an OR gate, as depicted in Fig. 13.12. The ZnS interference filters are operated in a bistable mode with a holding power of 20 to 40 mW and a fan-out of 3 to 4 in the experiment. Masking operations can be performed to block undesired non-overlapping portions. The 1×4 beams must be focused onto ZnS interference filters to achieve the intensity required for switching. A fly-eye lens array with 1-mm pixel separation is used for this purpose. Cascading, which has been demonstrated, is required for each channel.

Each channel has been constructed independently and shown to implement the sum rules correctly for all four input conditions. A lack of optical components has prevented simultaneous operation of all three channels.

It is noted that the Sagnac interferometer is an awkward way of dividing a beam into multiple beams. The combination of a phase grating or a hologram (discussed in Section 10) with an aberration-corrected and flat-field lens should produce an array of spots as close together as crosstalk will allow. One then can image the output from one filter onto the next filter. To take full advantage of free-space interconnection, the transmission peak of the ZnS IF must occur at the same wavelength everywhere on the filter, and the wavelength must be carefully chosen for perpendicular operation, with the filter situated normal to the beam.

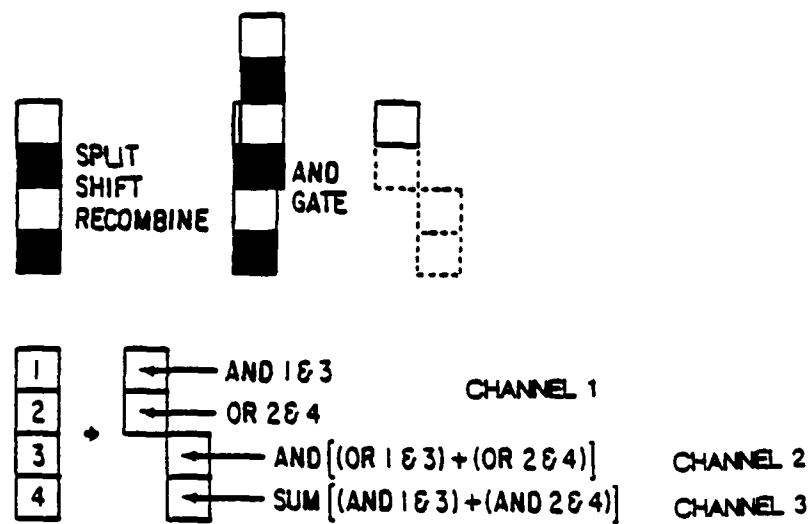


Figure 13.11 One manner of implementing the one-bit adder.

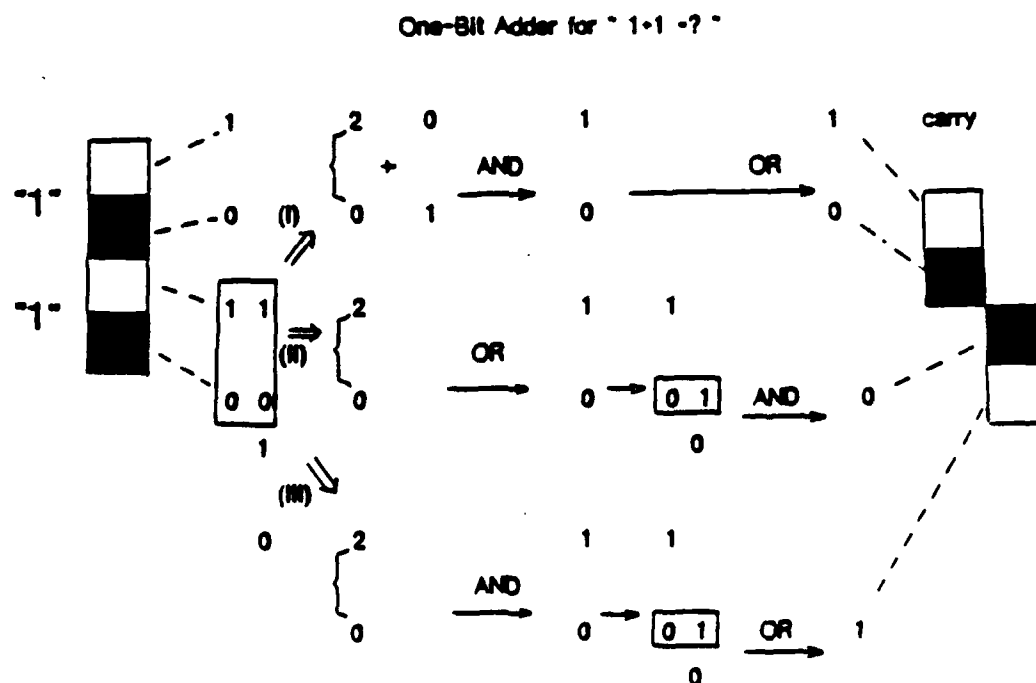


Figure 13.12 Explicit addition of "1 + 1" by a one-bit adder is shown schematically.

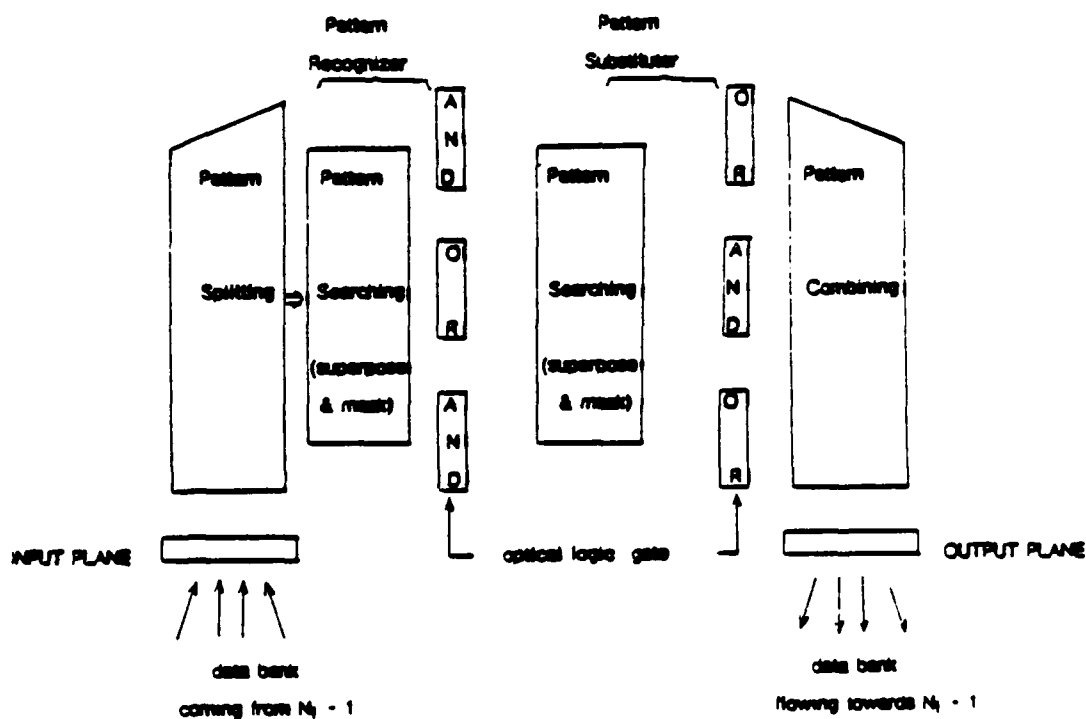


Figure 13.13 Block diagram of an n -bit adder.

13.5 Discussion

The advantages of optics, such as free-space interconnection and parallelism, and the use of 2-D arrays of bistable logic gates to make decisions, are clearly demonstrated in our simple symbolic substitution logic experiments. The extension of the current system to one having larger decision-making arrays with more stages, or even to one using faster optical logic gates such as GaAs to implement more sophisticated digital optical computing, will bring the following challenges.

First, although the dual-rail encoding of SSL simplifies thermal stabilization, doubling the device area results in some tradeoffs in the implementation of SSL. Another method using polarization encoding was proposed by K-H Brenner [13.3]. However, this method involves a more complicated optical system for encoding the input data and requires more optics for each iteration, leading to problems similar to those experienced in our experiments. In addition, combining input beams with

different polarizations is difficult. The optical system for SSL offers additional challenges to linear optics, including formation of several equal-power parallel beams from a single beam; splitting and recombining arrays of beams; and combining more than three inputs without loss.

Although SSL, in principle, could be applied to perform all arithmetic operations such as subtraction, multiplication, and division, it remains to be seen if SSL, in practice, will ever be implemented as a digital optical computer to compete with its electronic counterpart. The amount of optics required to perform an n -bit full adder will be substantial, based on the results of our one-bit adder experiment, and the problem of decreasing efficiency for each iteration must be addressed. This suggests that number-crunching is a questionable application for SSL in the future.

14. AN ALL-OPTICAL EXPERIMENTAL REALIZATION OF ASSOCIATIVE MEMORY [14.1]

As in Section 13, the all-optical systems described in this section demonstrate the use of nonlinear devices to make decisions. Conventional optical components are employed in the experiments described in Section 13 to manipulate the orientation of the processing beams. Here, in the associative-memory system, conventional optical techniques play the key processing role through matched filtering of a partial input. The nonlinear optical decision-making element utilizes the output of the matched filter, the cross-correlation beams, to determine the closest match. The corresponding hologram then is read with a counterpropagating beam.

14.1 Motivation and Concept

The experiments in Section 13 illustrate AND-gate operation, cascading, and a moderate amount of parallelism. They also illustrate the difficulties of dividing a beam into many beams so that each decision-making pixel receives the same amount of power in approximately the same cross sectional area. These geometrical optics difficulties and the high power requirements of large arrays of nonlinear decision-making pixels are formidable hurdles in the realization of massively parallel optical computing systems. Moreover, as seen from experiment, each two-spot recognition requires a path composed of a shifter, a lens system and a nonlinear decision-making element; extension of the technique of the three-spot pattern recognition to more complicated pattern recognition is improbable because the power continuously decreases and more optics are required for each subsequent path. It is concluded that, at this stage, it is not practical to perform pattern recognition in this way unless an innovative method to handle the regenerate feedback loop is devised. This conclusion suggests that the best application of nonlinear decision-making may not be in numerical computing, which is already accomplished easily with electronics. Rather, the best application may be to combine more traditional optical techniques of image processing, for example, real-time correlations, with a limited amount of nonlinear decision-making. Only if several correlation peaks or other events occur simultaneously will the system activate an appropriate response. An example of such an application is the associative memory.

An associative memory can be defined as a memory in which a stored piece of information can be retrieved as a whole when a probing input is only partially

complete or even distorted. Memory and thresholding functions are two essential ingredients to the associative memory, and the main advantage that optics offers is the capacity to provide the massive parallelism required for global interconnections. A Fourier hologram can be used to memorize or store a set of images, and the available recording materials range from emulsion-based to crystal-based materials. Decision-making feedback can be provided by the thresholding function of nonlinear devices such as ZnS or ZnSe IFs, GaAs etalons or photorefractive crystals.

Suppose one has a Fourier transform hologram formed by different object waves o_i and their corresponding reference waves r_i as shown in Fig. 14.1, where O_i and R_i represent the Fourier transforms of o_i and r_i , respectively. The hologram is recorded in the linear exposure region of a photosensitive recording material, and then the developed hologram returns to its original recording position, that is, $d_0 = d_1 = f$ as in Fig. 14.1. An incomplete or a distorted version \hat{O}_i of o_i is used as a reconstruction wave, therefore the amplitude transmitted by the developed hologram is proportional to

$$\begin{aligned} \hat{O}_i \sum_i |O_i + R_i|^2 &= \hat{O}_i \sum_i (|O_i|^2 + |R_i|^2) \\ &+ \sum_i \hat{O}_i O_i \bar{R}_i + \sum_i \hat{O}_i \bar{O}_i R_i + \text{intermodulation terms} , \end{aligned} \quad (14.1)$$

where \bar{O}_i and \bar{R}_i are the complex conjugate of transformed object waves O_i and reference waves R_i . The wave term $\sum_i \hat{O}_i \bar{O}_i R_i$ becomes $\sum_i (\hat{o}_i * \bar{o}_i) * r_i$ after one more Fourier transform operation, where $o_i \xrightarrow{\mathcal{F}} O_i$, $r_i \xrightarrow{\mathcal{F}} R_i$ and $*$ denotes a convolution operation. If \hat{O}_i is close enough to \bar{O}_i , then correlation \hat{o}_i and \bar{o}_i will have a sharp peak and r_i will be faithfully reconstructed. However, the faithful retrieval of the original object wave is of major concern in data processing. One way to retrieve o_i is to propagate the wave term $\sum_i \hat{O}_i \bar{O}_i R_i$ in Eq. (14.1) onto the hologram again, that is, to multiply by the stored wave term $\sum_i O_i \bar{R}_i$ in the hologram, to obtain the wave term $\sum_i [\hat{O}_i \bar{O}_i] |R_i|^2 O_i$. Therefore, o_i can be recalled

as long as \hat{O}_i and \bar{O}_i are closely matched [14.2]. An alternative to retrieving o_i is, first, to perform one more Fourier transformation and then to use nonlinear thresholding devices on the transform plane to perform decision-making, if R_i is a plane wave.

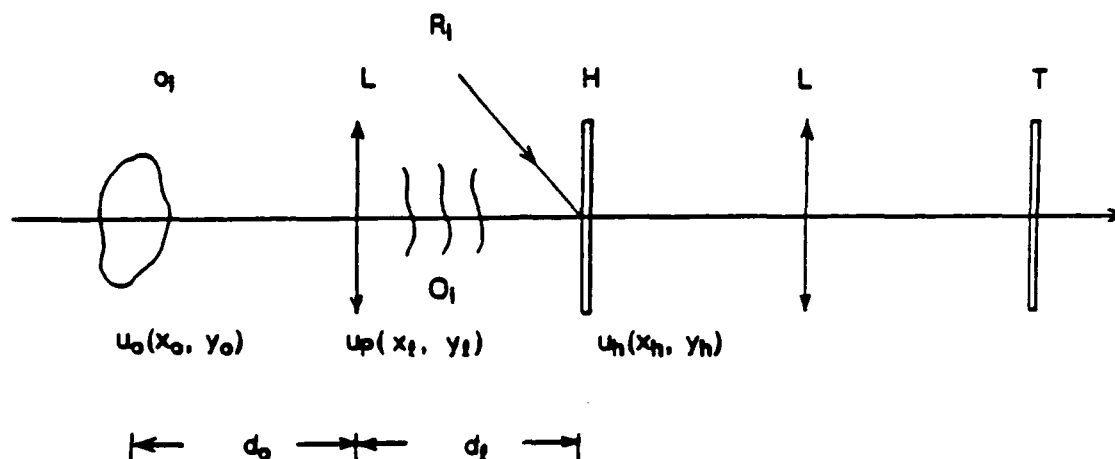


Figure 14.1 Schematic optical layout for recording a Fourier hologram and its application for associative memory.

Recent work involving associative memories based on various thresholding and feedback schemes [14.2, 14.4], and several optical implementations have been demonstrated [14.5-14.8]. The optical associative memory discussed here, which is similar to that presented by Soffer, et al., decides which stored object most closely matches a partial-object input [14.9]. The associative memory employs two elements. The first element, a Van der Lugt correlator, gives an analog 'degree-of-match' value, and the second, an optically bistable interference filter, gives a binary thresholding response.

The idealized system works by writing Fourier transform phase holograms, sequentially multiplexed in reference beam angle, of the different objects. The test object then is correlated with the stored objects, giving angularly separated cross-correlation beams. Each beam is focused onto an optically bistable IF where the decision of closest match is made at the central peak positions.

The interference filter has the following characteristics: structured $G(HL)^5 2H(LH)^5 G$ with a linear transmission of 54% at $\lambda = 513 \text{ nm}$; FWHM = 15 \AA ; and minimum switching power of 22 mW. The interference filter was used in conjunction with a holding beam in this experiment. Power is supplied, in this case

exactly counterpropagating with each correlation beam, in order to bias the device very close to the switching threshold. Only the minimum switching power then is necessary to push the device into the transmissive mode. This process reduces the required correlation power and thus the power through the hologram. The counterpropagating bias beam also facilitates the recall of stored objects.

One of the beams, corresponding to the most closely matched stored object, is sufficiently intense to provide the minimum switching power, allowing the holding beam to colinearly counterpropagate to the hologram. Because this beam is a plane wave at the proper incident angle to the hologram, it reads the appropriate object from the hologram, and the object is propagated back through the optical system. Also, because the hologram is thick, only this object is Bragg matched; no other objects are read out. This overall process is the auto-associative mode of operation. A partial object input recalls a reconstruction of the same object.

If the holding beam is propagated to another hologram or transparency a hetero-associative recall can also be performed. In this case any output can be associated with a matched partial input.

14.2 Experimental Apparatus

The experimental setup is shown in Fig. 14.2. In the recording stage, a collimated beam (the object beam) illuminates a transmissive object, a black-ink fingerprint on a slide. The object beam is transformed by a lens onto the Fourier plane where a holographic grating is formed by interfering with a collimated reference beam in dichromated gelatin (DCG). Two separate objects, labeled F2 and F3 (Fig. 14.3), were used in the experiment. F2 was interfered with a reference beam angle of 47° at the Fourier plane, and F3 was interfered with a reference beam angle of 58° .

Volume holograms have been shown to be better than the thin hologram in the application of matched filtering, resulting in better SNR, increased brightness of the correlation peak, and better spatial separation of the output signal [14.10]. DCG, one of the volume hologram materials, is a near-optimum hologram recording material, but it suffers from tedious processing. A considerable effort was made to optimize the DCG hologram [14.11]. More details can be found in Section 10.

The system require optimal discrimination to permit a minimal partial-object input. Switching power constraints for the optically bistable IF dictate some minimum auto-correlation efficiency. These requirements lead to a trade-off between system efficiency and discrimination. The object fingerprints, consisting of parallel curved ridges of roughly equal spacing, have a fairly low modulation depth, thus

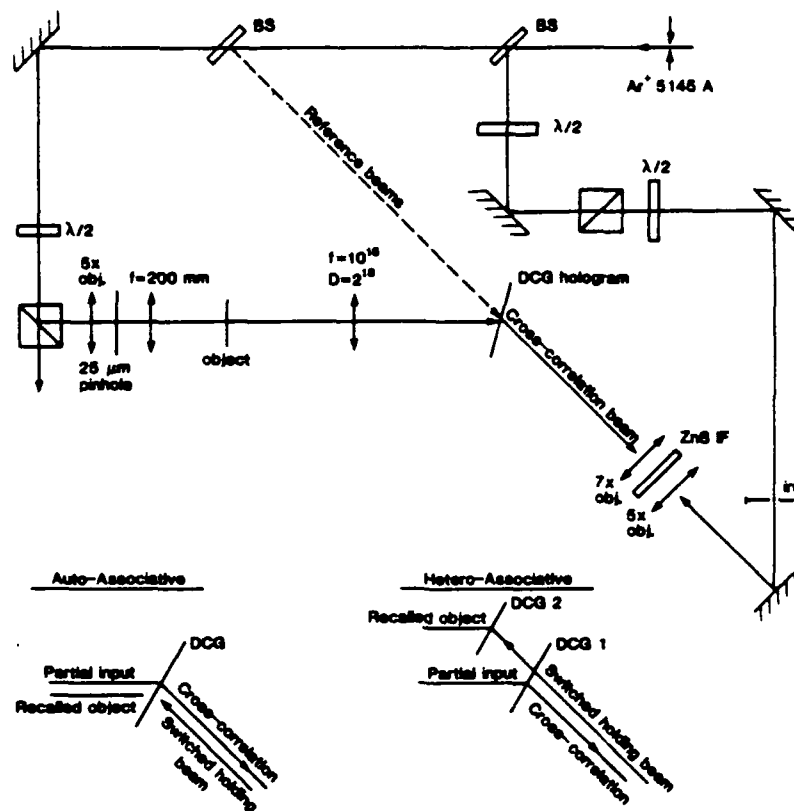
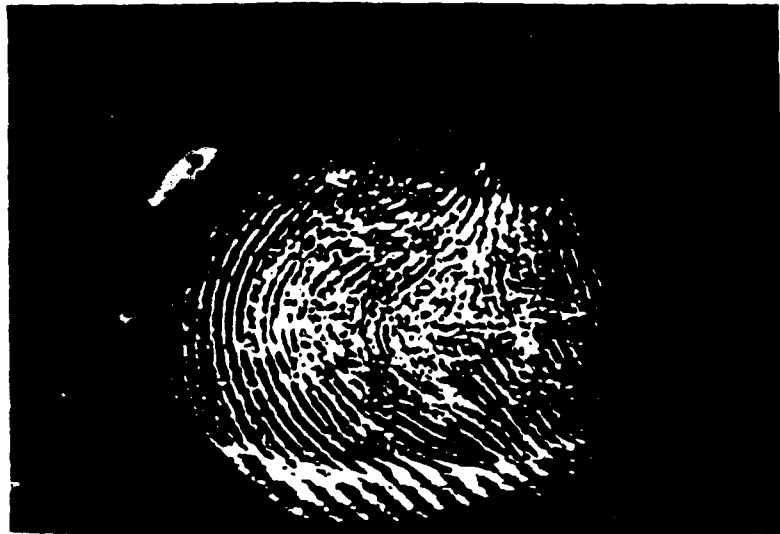


Figure 14.2 Experimental arrangement for associative memory.

only 20% of the power was in the diffracted "information ring" of the Fourier transform (Fig. 14.4). In addition, by line scanner, the relative intensities of the Fourier transform information showed a large dynamic range, from 1:15 to 1:500 compared to the zero-order. Such a large dynamic range generally cannot be accommodated by a recording film. This problem may be treated in two manners. First, one may wish to operate away from the Fourier plane, at a plane where the light distribution is more homogeneous. This should imply less object discrimination and less immunity to shift, scale and rotation. Second, one may wish to encode the object on a homogeneous spatial frequency "background," for example, a Lambertian plate. This method produces high discrimination due to the Lambertian alone and also has very poor scale, shift and rotation behavior.

(a)



(b)

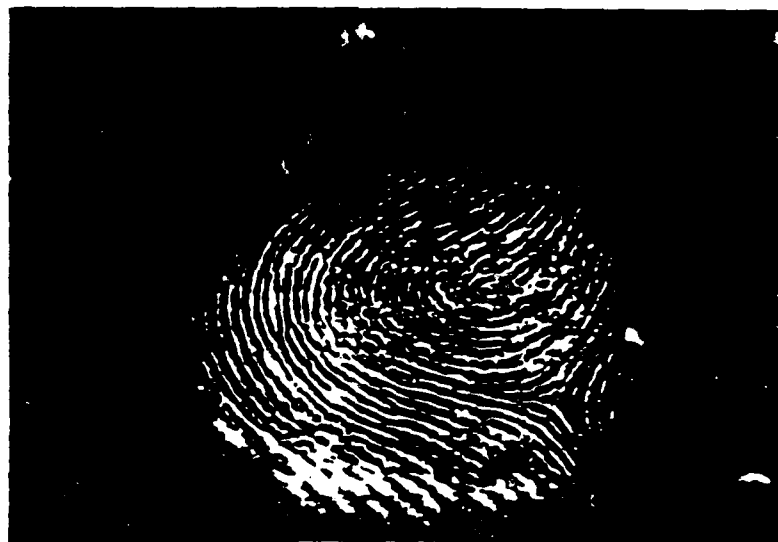
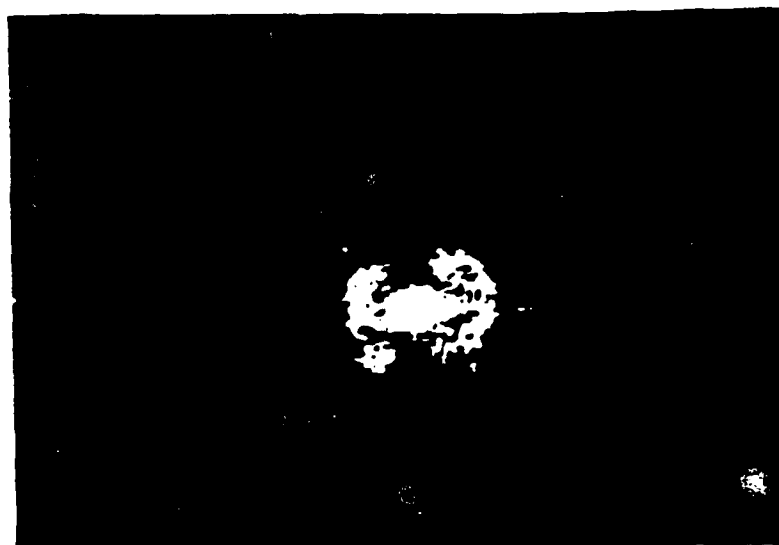


Figure 14.3 Object fingerprints: (a) F2, (b) F3.

(a)



(b)

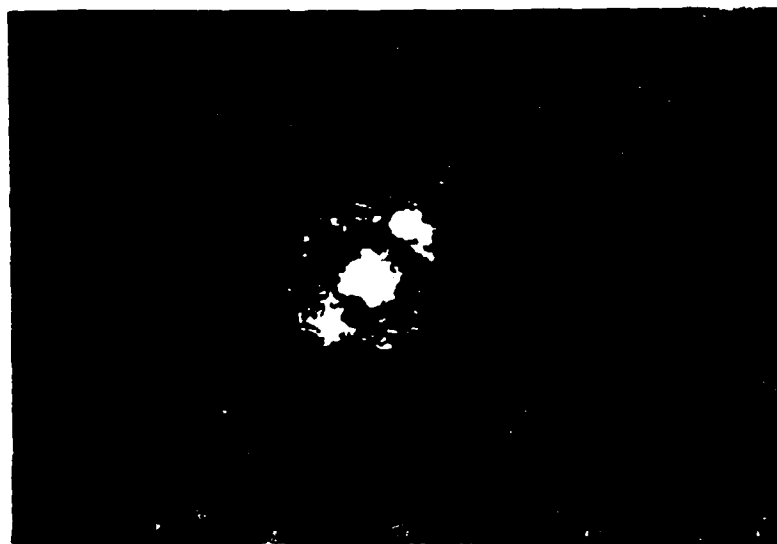


Figure 14.4 Photograph of the Fourier transform plane for the object fingerprints: (a) Fourier transform plane for F2; (b) Fourier transform plane for fingerprint F3.

Using carefully controlled objects, where the distinctive properties of the object will produce standard intensity Fourier transforms, diffraction efficiencies greater than 90% may be obtained. This efficiency degrades rapidly with the number of stored objects, depending on the method employed to store the holograms. The phase dynamic range of the holographic media is limited.

It is important to determine the appropriate reference/object beam ratio to give optimal grating visibility for the best efficiency while maintaining sufficient discrimination. It is also critical to ensure that the zero-order saturated the DCG, so that the holographic fringes are suppressed. This step eliminates the contribution to cross-correlation attributed to the zero order. Even when the zero-order saturates the DCG, some unwanted modulation occurs in the wings of the feature. Large diffraction from the zero-order would destroy the discrimination, because all objects have similar zero-orders. It was found that a reference-to-object ratio of 1:20 in the Fourier transform plane (taking the peak of the zero-order as the object intensity) and about 190 mJ/cm² exposure gave 4% auto-correlation efficiency; the efficiency-discrimination trade-off was in this way maximized.

The arrangement of relative spot-sizes between the holding beam and the switching beam is crucial in this experiment. It was observed that, under the well-aligned condition, a holding beam with a spot size larger than that of a switching beam is more effective in switching a bistable IF, rather than the opposite case. This is because the Gaussian-beam intensity profile of a holding beam has a smaller gradient for a large spot than for a small spot; the resultant intensity of the former case therefor would have a higher possibility of reaching the switch-on threshold. In addition, a better recall of the associative memory can be obtained if a higher contrast is available for the nonlinear etalon.

14.3 Results

During operation of the system, one of the objects was partially obstructed and illuminated (Fig. 14.5), causing auto-correlation beams to be diffracted from the hologram at the two reference angles. The autocorrelation peaks can be seen for F2 and F3 in Fig. 14.6, for the unobstructed and partial cases. For a partial input, the discrimination ratios (auto-correlation intensity versus cross-correlation intensity, in the correlation plane) of the correlation peaks were 50:1 with F2 as test object and 10:1 with F3. The reduced discrimination for partial F3 input was attributed to its smaller modulation depth, resulting in a larger zero-order content.

(a)

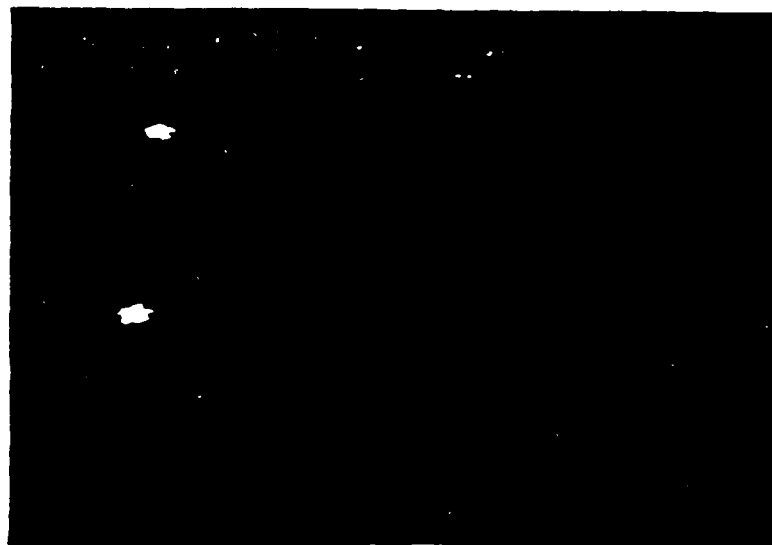


(b)



Figure 14.5 Partially obstructed objects: (a) F2, (b) F3.

(a)



(b)

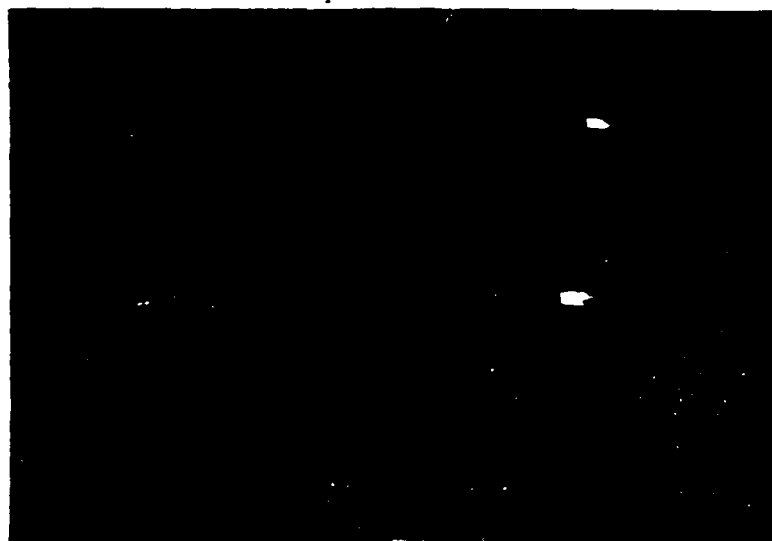


Figure 14.6 Cross-correlation beams for partial and whole input objects. (a) Upper; partial F2 input, Lower; whole F2 input. (b) Upper; partial F3 input, Lower; whole F3 input.

The auto-correlation peak for the partial-object input was of sufficient intensity (about 3 mW in a 100- μ m-diameter spot) to switch the optically bistable IF into the transmissive mode with a 20 mW holding beam (Fig. 14.7). Switching of the optically bistable IF allowed the counter-propagating holding beam to read out the hologram. The auto-associatively reconstructed image from a single readout beam (Fig. 14.8(a)) had a contrast ratio of 2:1. The hetero-associative reconstruction yielded similar results, as seen in Fig. 14.8(b). Because the diffraction efficiency of the hologram was only 4%, at least 75 mW of object input beam power was required, causing dynamic thermal damage in the DCG. The poor quality of the reconstructed objects was attributed to this thermal damage.

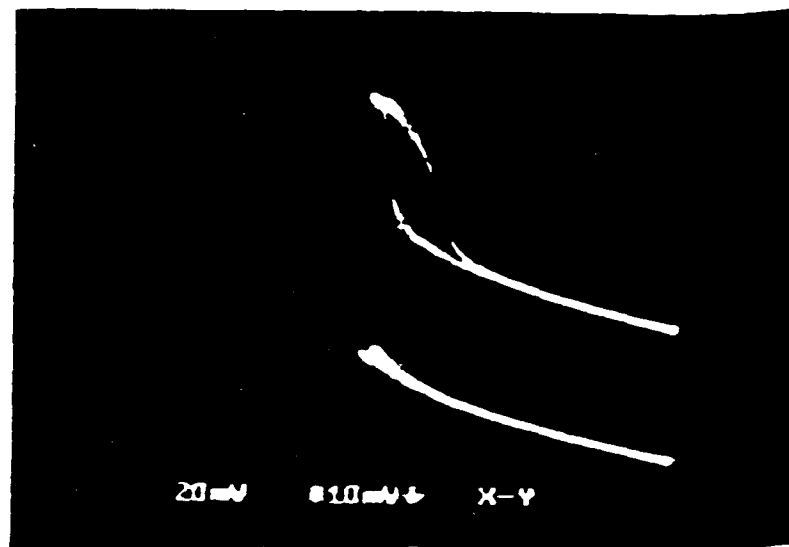


Figure 14.7 Optically bistable IF response: Upper, switched on by the correlation input (vertically off-set for clarity). Lower; switched off. (Trace describes input holding-beam intensity, $-x$, versus transmitted holding-beam intensity, y).

Decision-making was demonstrated on only one cross-correlation beam at a time, having determined that sufficient discrimination was present to preclude simultaneous switching of the beams. However, if two devices had been installed each would have produced an off-state recall (from the off-state linear transmission of the holding beam through the IF), further degrading the on-state recall discernibility.

(a)



(b)

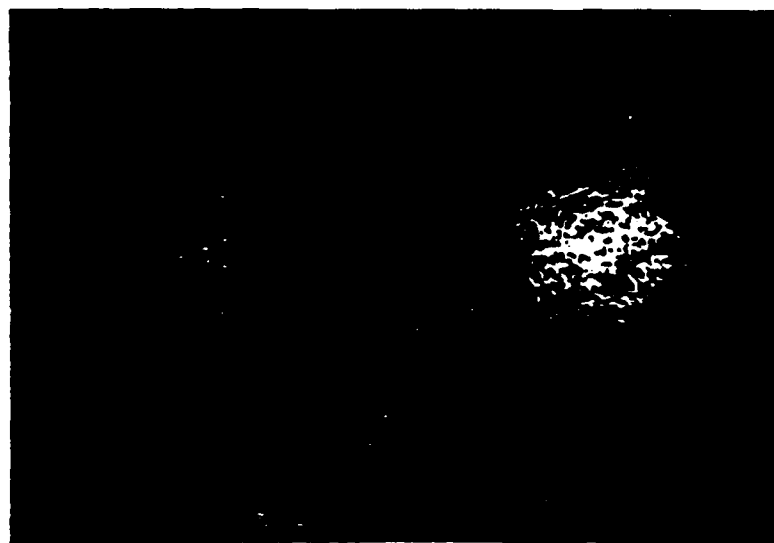


Figure 14.8 Recalled object F2. (a) Auto-associatively recalled: Left; switched off, Right; switched on. (b) Hetero-associatively recalled; Left; switched off, Right; switched on.

In a previous experiment one object was stored, the letters "WX," a partial input, was recognized, and associative recall was executed (Fig. 14.9). In this case discrimination was not necessary (only one object was stored) so that increased cross-correlation efficiency could be obtained through the use of the zero-order. High quality recall was facilitated because thermal damage was not an issue.

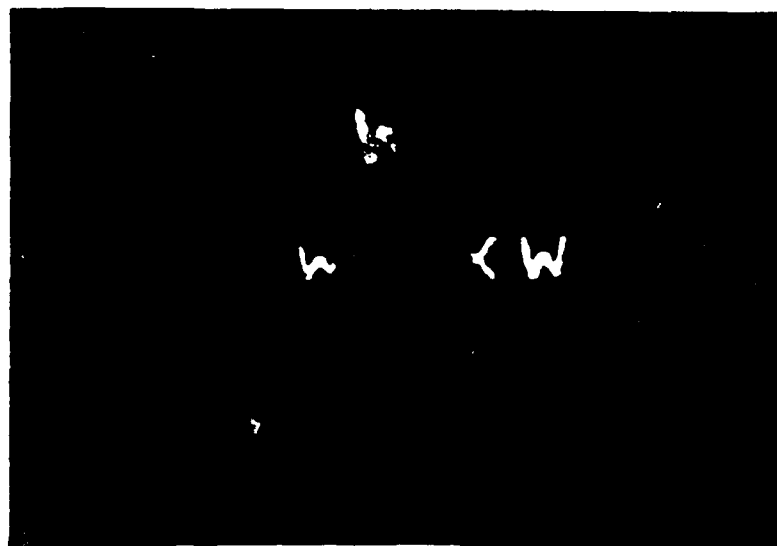


Figure 14.9 Recognition of partial input for stored "WX." Top; partial input, Bottom left; recall switched off, Bottom right; recall switched on.

14.4 Some Practical Details

It is well known that holographic recording materials have a relatively narrow linear range. Increasing the reference-to-object beam ratio will produce nearly linear recording, however, the diffraction efficiency strongly decreases. A tradeoff here can be obtained experimentally by plotting efficiency versus exposure with different reference-to-object beam ratios as parameters. Assume two exposing beams of incident intensities I_r and I_o , it has been shown that the diffraction efficiency η of real recording materials can be expressed as a function of the average exposure E_0 and visibility V [14.12]:

$$\sqrt{\eta(E_0, V)} = S(E_0, V) E_0 V \quad (14.2)$$

where $E_0 = (I_r + I_s)t$, t is the exposure time; $V \equiv \text{visibility} = 2\sqrt{I_r I_s} / (I_r + I_s) = 2\sqrt{(BR)/(1+BR)}$; $BR \equiv \text{beam ratio} = I_s/I_r$ or I_r/I_s ; and S is a non-constant factor. This procedure was performed to estimate the optimum exposure needed for recording the Fourier transform hologram in the associative memory experiment.

The Fourier transform of a fingerprint has a fairly large zero-order term, which places a large dynamic range requirement on the hologram to avoid distortions of the stored object information. These distortions come mainly from burnouts in the zero-order or dc term, which act as a high-pass filter, leading to edge enhancement that can sometimes help discrimination. However, this situation is not generally preferred, because any further adjustment of a lens after the occurrence of a burnout can encourage more burnouts to occur over the area that surrounds it. The information stored would be smeared or completely lost.

To prevent undesired burnout, a tolerance analysis of the positioning errors is made through an exact Fourier transformation in the following as one refers to the optical layout shown in Fig. 14.1. In addition, the deviation from plane wave illumination on an object, particularly the effect on the Fourier transform caused by spherical beam illumination, is discussed.

Let $u_o(x_o, y_o)$, $u_l(x_l, y_l)$, $u_h(x_h, y_h)$ be the electric field amplitudes on the plane of an object, a lens with focal length F and a hologram respectively. Each field has its own coordinate. Superscripts $-$, $+$ represent the field amplitude before and after that plane. From the diffraction theory [14.13], the field before the lens $u_l^-(x_l, y_l)$ can be expressed as follows.

$$u_l^- = \frac{e^{ikd_o}}{i\lambda d_o} q\left[x_l, y_l, \frac{1}{\lambda d_o}\right] \mathcal{F}\mathcal{F}\left\{u_o(x_o, y_o) q\left[x_o, y_o, \frac{1}{\lambda d_o}\right]\right\} \bigg|_{\substack{\xi \equiv \frac{x_l}{\lambda d_o} \\ \eta \equiv \frac{y_l}{\lambda d_o}}} \quad (14.3)$$

The field after the lens $u_l^+(x_l, y_l)$ is

$$u_l^+ = u_l^- e^{iknT} p(x_l, y_l) q^* \left[x_l, y_l, \frac{1}{\lambda F}\right] \quad (14.4)$$

where the quadratic phase factor is defined as $q(x, y, a) \equiv \exp[i\pi(x^2 + y^2)a]$; the

propagation constant $k = 2\pi/\lambda$; n is the refractive index of the lens material; e^{iknT} is a phase factor due to the lens thickness T ; and $p(x_l, y_l)$ is the pupil function of the lens.

Assume the lens is sufficiently large so that there is no vignetting, and with good homogeneity and constant thickness so that there is no phase error incurred when the waves propagate through it; therefore the e^{iknT} term can be dropped and $p(x_l, y_l)$ can be taken as unit over the range of integration. After some algebra, the diffracted field on the hologram $u_h(x_h, y_h)$ can be expressed as follows.

$$u_h = w \frac{e^{ik(d_o+d_l)}}{i\lambda d_o d_l} q\left[x_h, y_h, \frac{1-w/d_l}{\lambda d_l}\right] \cdot \iint dx_o dy_o u_o(x_o, y_o) q\left[x_o, y_o, \frac{1-w/d_o}{\lambda d_o}\right] e^{-ik(x_o x_h + y_o y_h) \frac{w}{d_o d_l}} \quad (14.5)$$

W is defined as $\frac{1}{w} \equiv \frac{1}{d_o} - \frac{1}{F} + \frac{1}{d_l}$, and $e^{ik(d_o+d_l)}$ results from the total phase shift along the optical path which is a constant and can be dropped.

Note that the Eq. (14.5) is an exact expression. Referring again to Fig. 14.1, four different conditions will be considered here according to practical situations: 1) a positioning error of the object; 2) a positioning error of the hologram; 3) an exact Fourier transform; and 4) a source of spherical illumination.

1) **A positioning error of the object, that is, $d_l = F$ but $d_o \neq F$.**

Let $d_o = F - \epsilon_o$, then we have

$$u_h = \frac{1}{i\lambda F} q\left[x_h, y_h, \frac{\epsilon_o}{\lambda F^2}\right] \cdot \iint dx_o dy_o u_o(x_o, y_o) e^{-i2\pi \left[x_o \left[\frac{x_h}{\lambda F} \right] + y_o \left[\frac{y_h}{\lambda F} \right] \right]} \cdot \left. \begin{array}{l} \xi \equiv \frac{x_h}{\lambda F} \\ \eta \equiv \frac{y_h}{\lambda F} \end{array} \right\} \quad (14.6)$$

The phase term $q(\dots)$ in the above expression indicates that the exact Fourier transform lies on a spheric surface of radius F^2/ϵ_o .

2) **A positioning error of the hologram, that is, $d_l \neq F$ but $d_o = F$.**

Let $d_l = F - \epsilon_h$ then we have

$$u_h = \frac{1}{i\lambda F} \iint dx_o dy_o u_o(x_o, y_o) q\left[x_o, y_o, \frac{\epsilon_h}{\lambda F^2}\right] e^{-i2\pi \left[x_o \left[\frac{x_h}{\lambda F} \right] + y_o \left[\frac{y_h}{\lambda F} \right] \right]} \quad (14.7)$$

To accurately represent the Fourier transform of an object, the phase factor $q(\dots)$ in the integrand of Eq. (14.7) should be equal to unity, that is, a condition $\epsilon_h < F^2 \lambda (x^2 + y^2)_{\max, l}^{-1}$ should be satisfied, where $(x^2 + y^2)_{\max, l}$ is the maximum extent of the field over a lens. This condition represents the positioning tolerance of a hologram that can be placed away from the exact Fourier transform plane to prevent a zero-order burnout, but still can maintain the fidelity of the Fourier transform. A numerical example is taken from the associative memory experiment: if $F = 254$ mm, $\lambda = 514.5$ nm and $(x^2 + y^2)_{\max} = 38$ mm, then $\epsilon_h = 870$ μ m.

Note that case 2) places a more stringent constraint on the positioning tolerance of Fourier transform than does case 1) if the intensity distribution rather than the field distribution is of main concern. This is because the quadratic term in case 1) is outside the integral, while in case 2) the quadratic term is inside the integral.

3) An exact Fourier transform, that is, $d_l = d_o = F$.

$$u_h = \frac{1}{i\lambda F} \mathcal{F}\mathcal{F}\{u_o(x_o, y_o)\} \left| \begin{array}{l} \xi \equiv \frac{x_h}{\lambda F} \\ \eta \equiv \frac{y_h}{\lambda F} \end{array} \right. \quad (14.8)$$

Note that the Fourier transform into the frequency domain is expressed in terms of $\left[\frac{x_h}{\lambda F}, \frac{y_h}{\lambda F} \right]$, which means that the scale of the distribution in the transform plane is dependent on the wavelength of the illumination and focal length of the lens.

4) A source of spherical illumination.

It is noted that an object, for example, a transparency, must be illuminated with uniform plane wave to give the exact representation of an object field $u_o(x_o, y_o)$ in Eq. (14.5). However, from a practical viewpoint, deviations from a real plane wave may occur. To some extent these deviations can be measured and corrected through optical testing techniques. The effect of the deviations can be simulated by a general situation, that is, through illumination by a spherical wave with a radius of curvature R . For a perfect plane wave, R equals infinity, a diverging-wave R is positive, and a converging-wave R is negative. Because a spherical wave can be represented by a

quadratic expression, $u_o(x_o, y_o)$ in Eq. (14.5) is replaced by $\left[\frac{1}{i\lambda R} q\left(x_o, y_o, \frac{1}{R\lambda}\right) \right] u_o(x_o, y_o)$. The phase factor $q(\dots)$ in the integrand becomes $\frac{1}{i\lambda R} q\left(x_o, y_o, \left[\frac{1}{d_o} - \frac{w}{d_o^2} + \frac{1}{R}\right]/\lambda\right)$ after taking into account the spherical wave illumination. Other terms in Eq. (14.5) remain the same.

Therefore, the Fourier transform can be obtained only when

$\frac{1}{d_o} - \frac{w}{d_o^2} + \frac{1}{R} = 0$, that is, when

$$\frac{1}{d_f} = \frac{1}{F} - \frac{1}{R+d_o} \quad (14.6)$$

This expression indicates that under a spherical illumination the position of the Fourier transform depends on d_o and R . When $R = \infty$, that is, under plane wave illumination, the Fourier transform is still at the plane $d_f = F$. In practice, the choice of a collimating lens is important. In general, a lens of longer focal length has a larger radius of curvature, which can tolerate more collimation errors. A larger aperture, however, is required to prevent vignetting. In contrast, a lens of shorter focal length has less tolerance of the collimation error. Use of a shorter lens may cause undesired diffraction, incurred by the correspondingly small aperture, which affects the efficiency and resolution of the Fourier transform hologram.

14.5 Summary and Discussion

An associative memory for fingerprint identification has been constructed using a Van der Lugt correlator and an IF as a thresholding device.

In the associative memory the DCG holographic Van der Lugt filter limited the performance of the system. Although the object employed was responsible for certain aspects of the experimental difficulty, even under ideal object conditions the diffraction efficiency achievable will be a limiting factor. This problem is aggravated as the number of stored objects is increased [14.14]. Photorefractive materials may serve to increase the number of stored objects, by means of their large material thicknesses, but typically the diffraction efficiencies are even more limited than with DCG holograms (see Section 12). The power throughput requirements dictated by subsequent nonlinear devices can also introduce serious complications, because of the storage and erasure dynamics of photorefractive holograms. These dynamics can also introduce new flexibility to the holographic element.

Optical nonlinear devices were successfully employed as decision-making elements in these all-optical systems. However, the limiting aspects of the nonlinear devices and the holographic element were felt even in these comparatively simple systems.

15. ALL-OPTICAL COMPARE-AND-EXCHANGE SWITCHES

15.1 Introduction

All-optical compare and exchange is experimentally demonstrated using ZnS optical bistable devices. The compare-and-exchange demonstration utilizes polarization multiplexing and filtering, latching, and bidirectional logic. The combination of two-dimensional arrays of compare-and-exchange modules with optical perfect-shuffle interconnections leads to pipelined optical sorting networks that can process large numbers of high-bandwidth signals in parallel. Optical sorting networks with these characteristics are applicable in telecommunication switches, parallel processor interconnections, and database machines.

Sorting is one of the most common and well-understood topics in computer science. Serial-sorting algorithms require at least $O(N \log N)$ temporal complexity [15.1]. Hardware based on parallel sorting algorithms offers enhanced performance on problems that must rapidly sort large quantities of information. Since the number of clock cycles, devices, and interconnects are limited resources in any processing environment, we need parallel algorithms with sublinear temporal and practical spatial complexity. In addition, the algorithms we choose must be optimum with respect to our specific implementation technology. For instance, the mesh algorithms developed for VLSI require only nearest-neighbor connections and are sublinear $O(N^{1/2})$ in temporal complexity [15.2]. Unfortunately, mesh algorithms must finish sorting one sequence before beginning another; thus, their throughput is limited by their latency. Although the shared memory [15.3] and some network [15.4] algorithms have the lowest temporal complexity $O(\log N)$ of all sorting algorithms, but are not practical with current technology because they require globally reconfigurable interconnects and excessive spatial resources, respectively.

Network algorithms based on the bitonic sort [15.5] have sublinear temporal complexity $O(\log^2 N)$. Moreover, they can be pipelined in stages for high throughput; and thus are useful in problems where throughput is as critical as latency. But the bitonic sorting network requires at least one globally connected interstage communication pattern. For instance, the perfect-shuffle [15.6] connection pattern transmits half the information present in the top half of a list to the bottom half and vice versa. Because VLSI is confined to the two-dimensional surface of a chip and electrons in wires are capacitively coupled, practical electronic perfect shuffles are limited to small numbers of channels and low data rates. In contrast, the noninteracting nature of photons and three-dimensional connection capability of optics

allows optical perfect-shuffle networks to have large numbers of parallel channels and high data rates [15.7-15.9]. Thus, optical sorting networks based on the perfect-shuffle interconnection and the bitonic algorithm are desirable when the number of communication channels or the data rates exceed the capabilities of electronic systems.

In particular, optical sorting networks are applicable in telecommunication switches that route high-bandwidth optical data packets [15.10]. Telecommunication switches must handle many parallel channels, have low latency, and keep up with the packet-generation rates. Similarly, high-throughput sorters serve as the communication fabric of electronic multiprocessors [15.11]. In these parallel processors, the number of processing elements, and thus the computational power, is governed by the number of parallel data channels. Furthermore, the throughput of each processing element is limited by the latency and throughput of each interconnection. In addition, sorting hardware may serve as dedicated subsystems for parallel database operations [15.12] in conjunction with optical memories [15.13]. Parallel and independent memory access can generate data rates beyond the capabilities of electronic systems.

Network sorting algorithms need, in addition to perfect-shuffle interconnections, 2×2 self-routing crossbar switches where each routing decision depends on the relative magnitude of the local information. Hence, we desire implementations of the 2×2 self-routing crossbars that are compatible with optical interstage connections and fulfill the requirements of bandwidth and parallelism in sorting applications. The function of such self-routing crossbars can be separated into the operations of comparison and exchange: comparison determines the relative magnitude of the local data; exchange configures the crossbar switch dependent on the outcome of the comparison.

In all subsequent discussions, we assume a binary representation for the data. Overscores represent the invert operation; thus, \bar{R}_1 and \bar{R}_2 are the logical complements of the system reset. Brackets "[]" contain a latching condition which we will explain shortly.

An algorithm for compare and exchange proceeds as follows: we label the synchronous input channels A and B, and operate serially from most to least significant bit. If $A_i > B_i$ occurs before $B_i > A_i$, where i represents the bit position, then the switch latches into the "don't-exchange" position. Conversely, if $B_i > A_i$ occurs, then the switch latches into the exchange position (see Fig. 15.1). Latching implies that once an inequality has been detected, the exchange switch becomes set into one particular configuration until the system is reset.

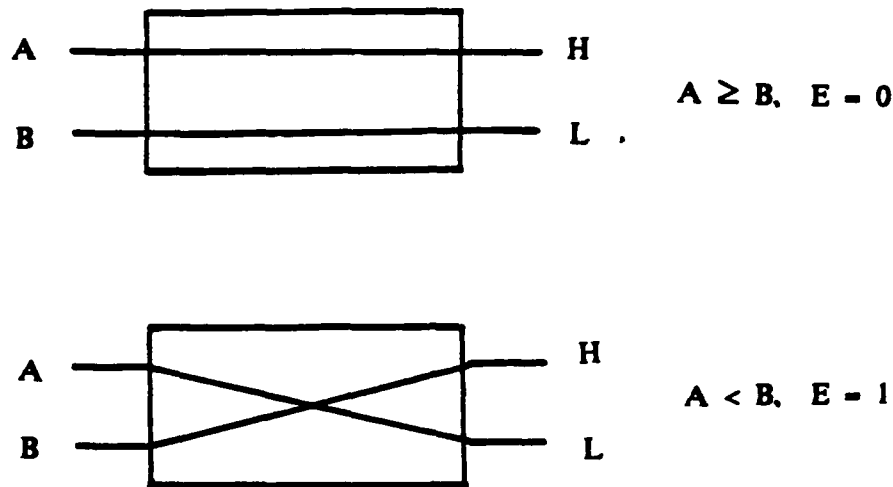


Figure 15.1 The compare-and-exchange module. E represents the exchange signal, A, B represent the two input numbers, and H, L represent the higher number and the lower number, respectively.

Optical bistable devices have the potential for high-speed optical signal processing and computing [15.14, 15.15]. ZnS and ZnSe bistable interference filters have already been used to demonstrate simple digital optical circuits, pattern recognition, symbolic substitution, and one-bit addition, because they can be operated in the visible spectrum and are relatively easy to fabricate [15.16, 15.17]. We have experimentally demonstrated a circuit that performs compare and exchange with ZnS interference filters as bistable devices. Here the ZnS interference filters are used in less common modes of operation including latching and bidirectional logic. In addition, we have employed polarization multiplexing and filtering to achieve channel isolation, four-port bidirectional devices, and reduced feedback.

15.2 Compare-and-Exchange Circuit Design

More than one circuit design is possible for comparison [15.18]. Figure 15.2 is a schematic of our comparison circuit and consists of three parts. The first part is a comparator to distinguish between the cases where $A_i > B_i$ or $A_i < B_i$ (Fig. 15.3 illustrates this by generating $A_i \bar{B}_i$ and $\bar{A}_i B_i$). In the second and third parts, two latching gates are employed, so that when $A < B$, $\bar{A}_i B_i = 1$ comes first, and one latching gate will be switched on to give an exchange signal $E = 1$. The gate remains in the on state until all the bits of A and B are transmitted. Similarly, when $A > B$, $A_i \bar{B}_i = 1$ comes first, and another latching gate will be switched on to

prevent the exchange. From Fig. 15.3 we see that $A_i \bar{B}_i = 1$ and $\bar{A}_i B_i = 1$ never occur simultaneously, making it possible to separate the state of the latching gates.

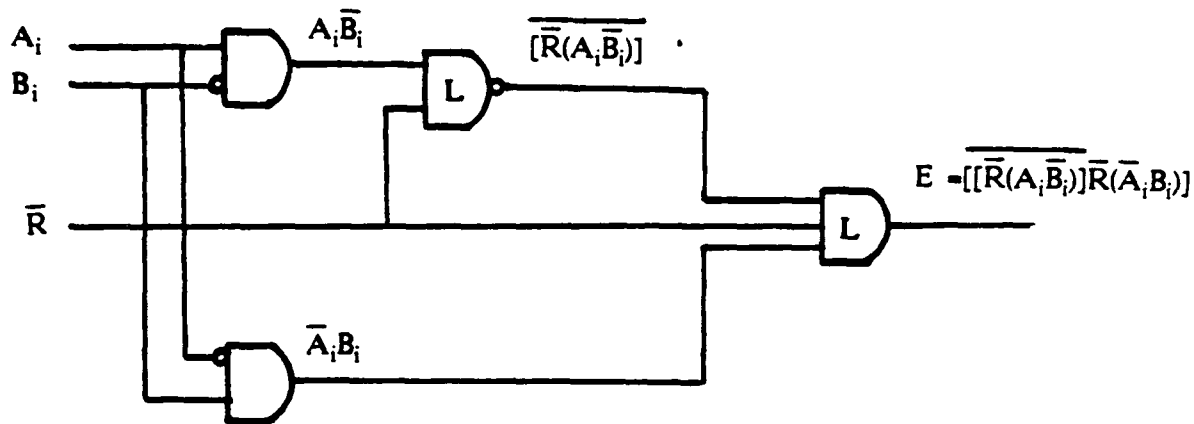


Figure 15.2 Compare circuit in which *L* indicates the latching logic gate.

A_i	B_i	$A_i \bar{B}_i$	$\bar{A}_i B_i$
0	0	0	0
0	1	0	1
1	0	1	0
1	1	0	0

Figure 15.3 The use of $A_i \bar{B}_i$ and $\bar{A}_i B_i$ to compare A and B . When $A \cup B$, $A_i \bar{B}_i = 1$ will appear first. When $A \cap B$, $\bar{A}_i B_i$ will appear first.

However, there is more than one way to implement exchange. The appropriate choice depends on the application requirements, the technology characteristics, and the corresponding comparison circuit. For demonstration purposes, we will construct an active exchange module. If the exchange signal $E = 1$ is present, it sends B to the H output and A to the L output. Otherwise, if the exchange signal is 0, it sends A to the H output and B to the L output.

The above discussion shows that our circuit design needs a comparator, two latching gates, and an exchanger. Figure 15.4 shows that $A_i \bar{B}_i$ and $\bar{A}_i B_i$ can be generated from a single bistable etalon by using its reflections from both sides, such that a single bistable device can be used as a comparator. Also, the latching

operation is natural for bistable devices, so that the two latching gates are just two bistable devices. Another bistable device is used as the exchanger with its transmission determined by the exchange signal. The details of their operations are discussed in the following section.

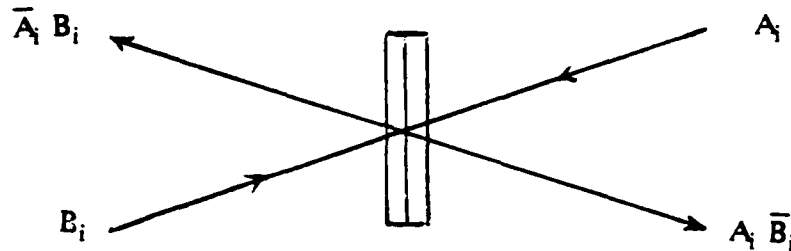


Figure 15.4 Generation of $A_i \bar{B}_i$ and $\bar{A}_i B_i$ using a single bidirectional, reflection-mode Fabry-Perot etalon.

15.3 Experimental Demonstration

The experimental layout for the compare-and-exchange circuit is shown in Fig. 15.5. An argon laser and a phase grating generate four optical beams, each having a peak power of about 40 mW. A chopper modulates beams A and B with the test sets, and blocks the holding beams \bar{R}_1 and \bar{R}_2 between each test set to allow the latching gates to reset. A half-wave plate gives the two holding beams \bar{R}_1 and \bar{R}_2 s polarization.

For the experimental demonstration of all-optical compare and exchange, we choose test vectors of $A = 110001$, $B = 101011$ and $A = 100011$, $B = 110101$. In the former set of test vectors, $A_i > B_i$ occurs first; permutations of $A_i B_i$ follow to ensure that the switch is properly latched. Similarly, for the latter group of test vectors we find that $B > A$ and demonstrate the exchange stability to further permutations. Until the first mismatch, the position of the exchange switch is not important to first-order approximation, because the output data streams are identical. In Fig. 15.6, we depict the expected operation of the latching compare and passive circuits described above for both test sets. All data used in the simulations are based on the structure of each filter. The curves are drawn upside down to be consistent with the experimental photographs. We see that whether $A > B$ or $A < B$, the larger number always goes to the H output. We did not fit the simulations to the experimental

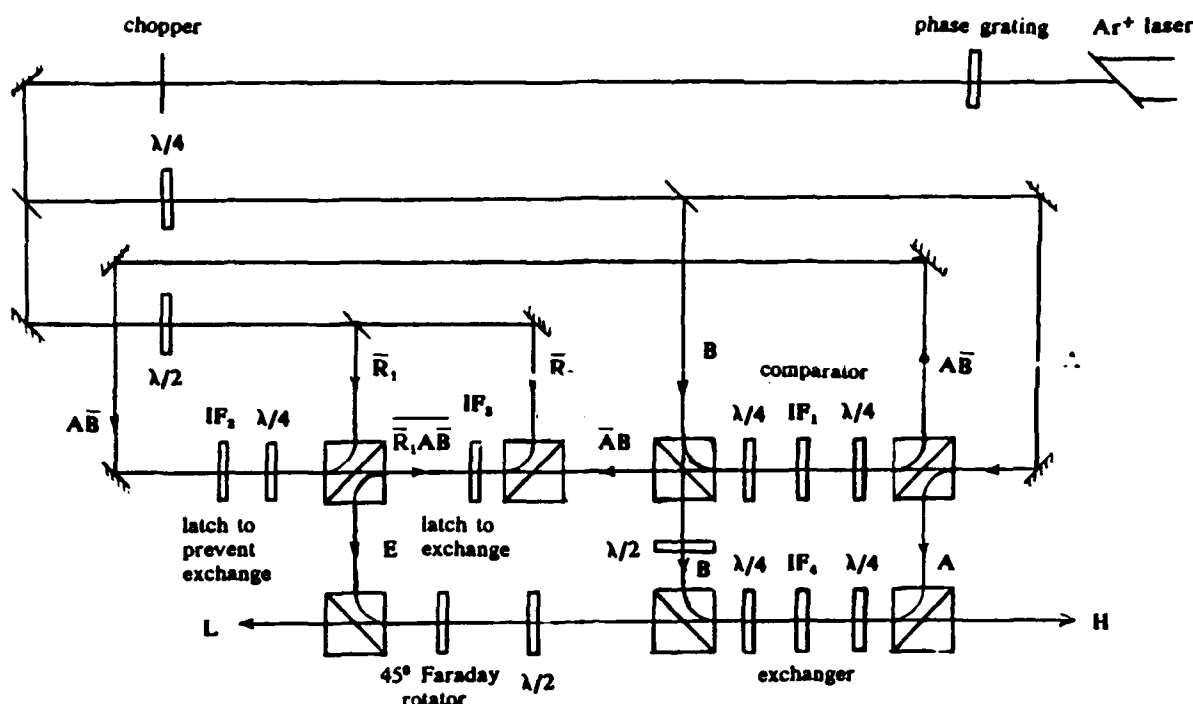


Figure 15.5 Experimental layout for all-optical compare and exchange with IF-ZnS interference filter, $\lambda/2$ - half-wave plate, $\lambda/4$ - quarter-wave plate. E represents the exchange signal, A , B represent the two binary encoded numbers, and H , L represent the outputs of the larger number and the smaller number, respectively.

results because we wanted to show the ideal results with suitable devices. The transfer functions shown in Fig. 15.7 used the same data.

The compare circuit operates in the following manner. The circularly polarized data beams, A and B , are incident on two polarizing beam splitters (PBSs). These PBSs serve two functions. One function is to sample the data beams for the compare operation: the p polarization from the A channel propagates through the PBS for comparison, the s polarization is reflected to the exchange switch: conversely, the s polarization of the B beam is reflected for comparison and the p polarization propagates through the PBS to the exchange switch while its polarization is rotated by the half-wave plate to match that of A . The orthogonally polarized data beams that were injected into the compare circuit are converted to circular polarization by two quarter-wave plates. The circularly polarized data beams are incident on the first interference filter (IF_1).

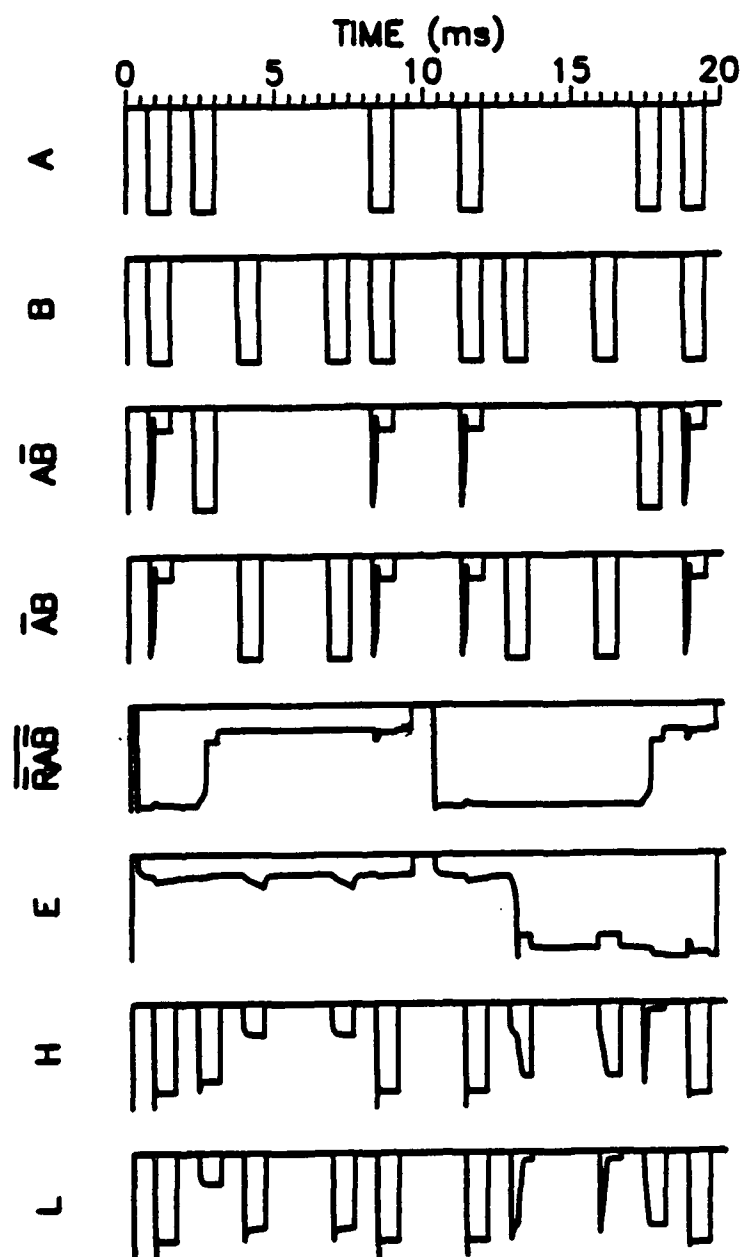


Figure 15.6 Expected operation of compare-and-exchange circuits test set 1 and 2 on horizontal axis. All curves are drawn upside down consistent with the experimental photographs. The vertical axes are in arbitrary units.

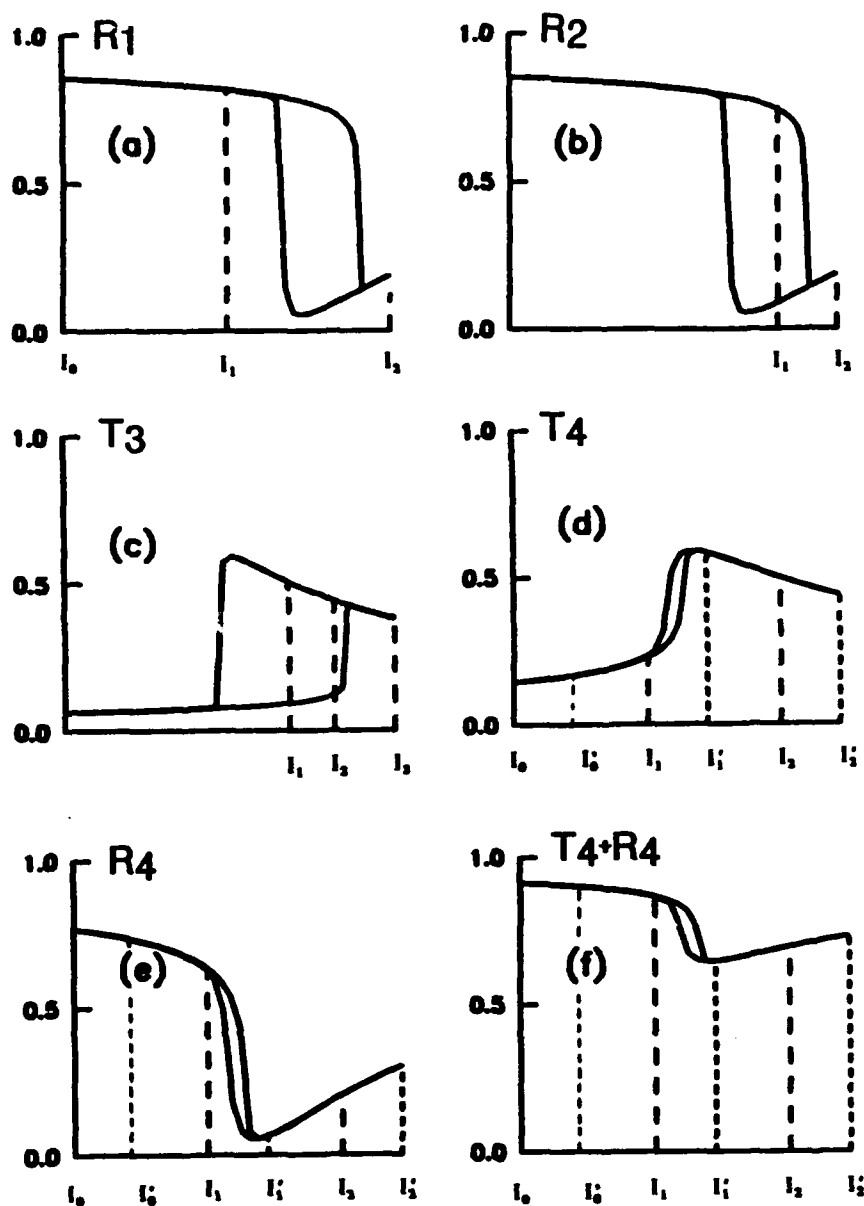


Figure 15.7 Computer-simulated transfer functions for the ideal interference filters. The horizontal axes are input power in arbitrary units and the vertical axes are reflectivity or transmissivity of the filters. (a) Reflectivity R_1 of the comparator IF_1 with $I_0 = 0$, $I_1 = A_i$ (or B_i), and $I_2 = A_i + B_i$. (b) Reflectivity R_2 of the latching NAND gate IF_2 with $I_1 = R_1$, and $I_2 = R_1 + A_i B_i$. (c) Transmissivity T_3 of the latching AND gate with $I_1 = R_2$, $I_2 = R_2$, $I_3 = R_2 + R_1 A_i B_i$ (or $R_2 + A_i B_i$), and $I_4 = R_2 + R_1 A_i B_i + A_i B_i$. (d) Transmissivity T_4 of IF_4 with $I_0 = 0$, $I_1 = E$, $I_2 = A_i$ (or B_i), $I_3 = A_i + E$ (or $B_i + E$), $I_4 = A_i + B_i$, and $I_5 = A_i + B_i + E$. (e) Reflectivity R_4 of IF_4 with definitions as in (d). (f) The sum of (d) and (e).

IF_1 operates in reflection mode as a bidirectional comparator. If B_i is zero and if A_i is one, then the circularly polarized A_i is reflected by IF_1 , converted to the s polarization by the quarter-wave plate, and reflected by the PBS to produce the signal $A_i\bar{B}_i$. In a similar fashion, if A_i is zero and B_i is one, then the circularly polarized B_i is reflected by IF_1 , converted to the p polarization by the quarter-wave plate, and transmitted through the PBS to produce the signal \bar{A}_iB_i . Thus, the PBSs also function as part of a bidirectional switch. Because the filter inputs are A_i and B_i and the outputs are both \bar{A}_iB_i and $A_i\bar{B}_i$, IF_1 is a four-port device.

$[\bar{R}_1A\bar{B}]$ is the exchange-prohibited signal from IF_2 , which works in reflection mode as a latching NAND gate, as shown in Fig. 15.7(b). As long as $A_i\bar{B}_i = 0$, the reflection of \bar{R}_1 is high, as it has been polarization rotated. It passes through the PBS, causing IF_3 to switch on and the latch to have a high transmission when \bar{A}_iB_i becomes 1, as seen in Fig. 15.7(c). This is the exchange situation with $E = 1$. If $A_i\bar{B}_i$ becomes 1, IF_2 switches on and latches to have a low reflectivity; the reflection of \bar{R}_1 will be low thereafter. If this sequence takes place before the first occurrence of \bar{A}_iB_i equal to 1, IF_3 will never achieve enough input power to switch on, and the output of exchange signal E will always be low.

The final filter IF_4 is the exchanger, and works in both transmission and reflection modes, as shown in Figs. 15.7(d) through (f). If both A_i and B_i are 0, both outputs are 0 independent of the exchange signal. In the case that $E = 0$ and only one of A_i and B_i is 1, IF_4 will not switch on; beam A_i reflects to the high output on the right; beam B_i reflects to the low output on the left. If both A_i and B_i are 1, IF_4 switches on and has a high transmissivity and low reflectivity. Both sides have a high transmission independent of the exchange status. The exchange control signal E will move the transmission curve closer to the laser frequency. When $E = 1$, either signal (or both) can switch on the gate. A_i and B_i then are transmitted to the opposite sides, in other words, they are exchanged.

Figure 15.8 shows the results of the comparator. It demonstrates clearly that as soon as there is a difference between A_i and B_i , the comparator has a high output to the following corresponding gate, which then makes the appropriate decision. At the first bit, numbers A_i and B_i are equal. The output changes rapidly from a high reflection to a low reflection and produces a sharp peak pulse at the rising edge of the output. If the signal pulse width is large enough compared to the width of the sharp pulse, this sharp pulse will not have enough power and would not switch the next stage. Before each comparison of the input numbers, \bar{R}_1 and \bar{R}_2 are reset to 1. When the compare and exchange is complete, \bar{R}_1 and \bar{R}_2 are shut off. The system waits for the next operation. Figure 15.9 shows the exchange-prohibited signal

$[\bar{R}_1 A \bar{B}]$. When $A_i > B_i$, this signal latches to a low output. Figure 15.10 shows the exchange signal E which is the transmission of \bar{R}_2 . On the left, although two cases of $B_i > A_i$ occur, E remains in its low state because the earlier occurrence of $A_i > B_i$ latched the exchange-prohibited signal to 0. On the right, A is larger than B . When $A_i > B_i$ first occurs, E is latched to 1, and therefore all the following bits exchange their positions. The time delay at the rising edge of E is caused by the switching speed of the device. Figures 15.11 and 15.12 show the results of compare and exchange in the two cases of $A > B$ and $A < B$, respectively.

15.4 Discussion

From the experimental results of the all-optical compare-and-exchange circuit, one can see that the contrasts are not as good as those in the simulations. This is because the filters used in the experiment were not specifically designed for reflection-mode operation. Therefore, the low state of the reflection is higher than we expected. However, even with such nonoptimal filters, the system worked. Specifically designed reflection-mode filters, would improve the contrasts of the outputs and decrease the power required. Another bistable optical device such as a bistable laser amplifier, with its threshold set half way between the worst case levels 0 and 1, could amplify the outputs of the exchanger as well as enhance the contrast. Then the outputs could be used to drive the next compare-and-exchange module in a self-routing optical network.

It was difficult to obtain stable simultaneous operation of all four interference filters long enough to test the system, especially since the contrasts of the devices are not good. The data in Figs. 15.8 through 15.12 were taken with only the relevant section working. Figures 15.8 through 15.10 were taken from the compare unit consisting of IF_1 through IF_3 , and Figs. 15.11 and 15.12 were the results from the exchange unit of IF_4 . While performing the exchange, we used a third beam having the power consistent with the exchange control signal E . Therefore, the experimental results do not show the delay as seen in the simulations. We also performed the experiment with IF_1 and IF_3 , producing a real exchange signal for the last gate, to show that when $A < B$, the exchanger works.

It also was not easy to focus A , B , and \bar{E} onto IF_4 and yield H and L without energy losses when the respective polarization directions were considered. A 45° Faraday rotation glass and a half-wave plate placed between the exchange control signal E and IF_4 might solve this problem. A plane-polarized light beam passing through the glass will have its polarization direction rotated through an angle θ relative to the polarization direction of the incident beam. A beam coming from the

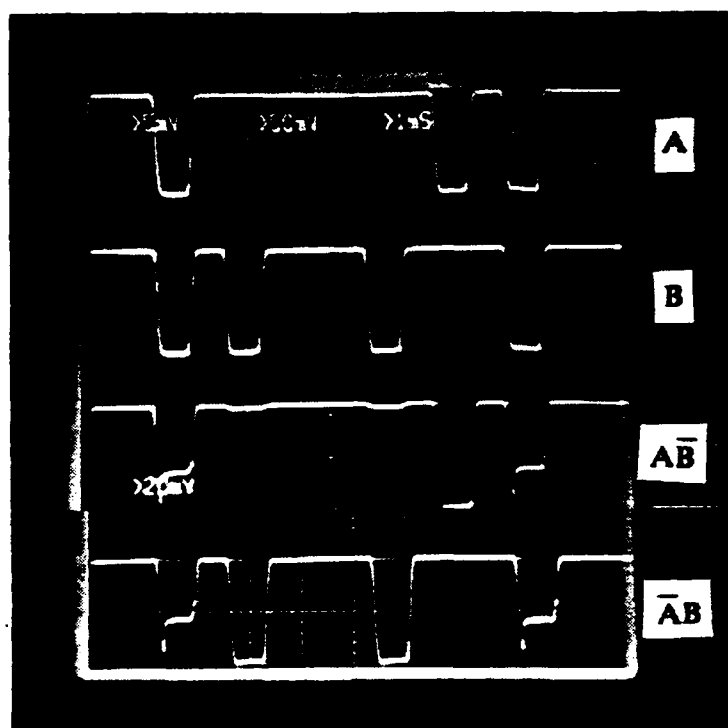


Figure 15.8 Experimental results of the inputs A , B and the logic outputs AB , $\bar{A}\bar{B}$. The input powers are 11 mW each, and the output power is about 5 mW.

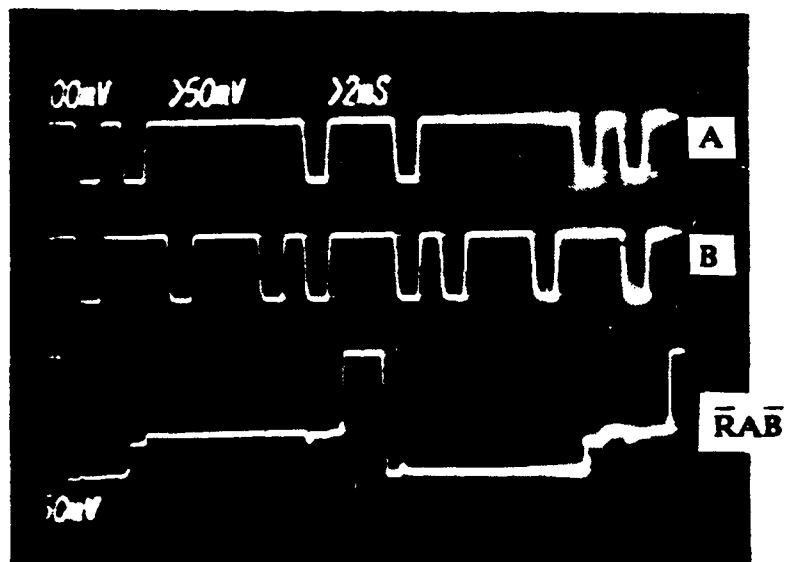


Figure 15.9 Output of the exchange-prohibited signal $\bar{R}_1 \bar{A} \bar{B}$. The upper two traces show the two groups of numbers coming into the system. In the first group, A is larger than B and in the second one, B is larger than A . The power of the holding beam is 19 mW, and the output power is about 6 mW.

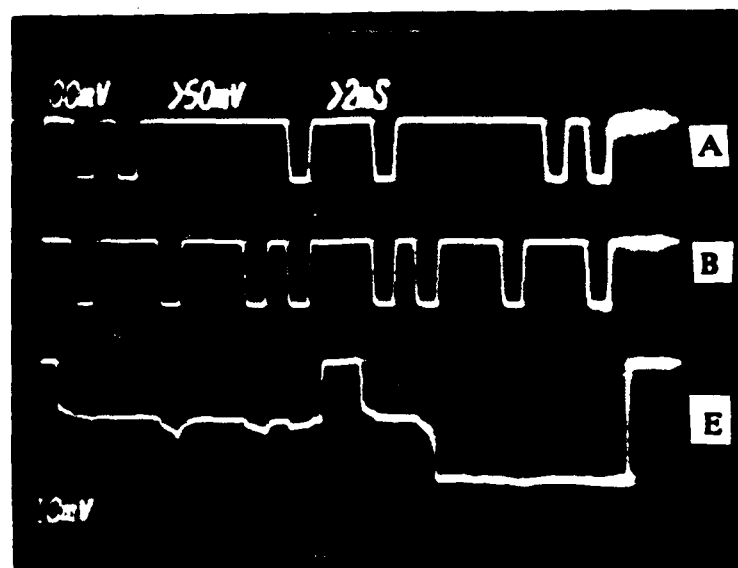


Figure 15.10 Output of the exchange signal E. The holding power is 20 mW and the output power is 5 mW.

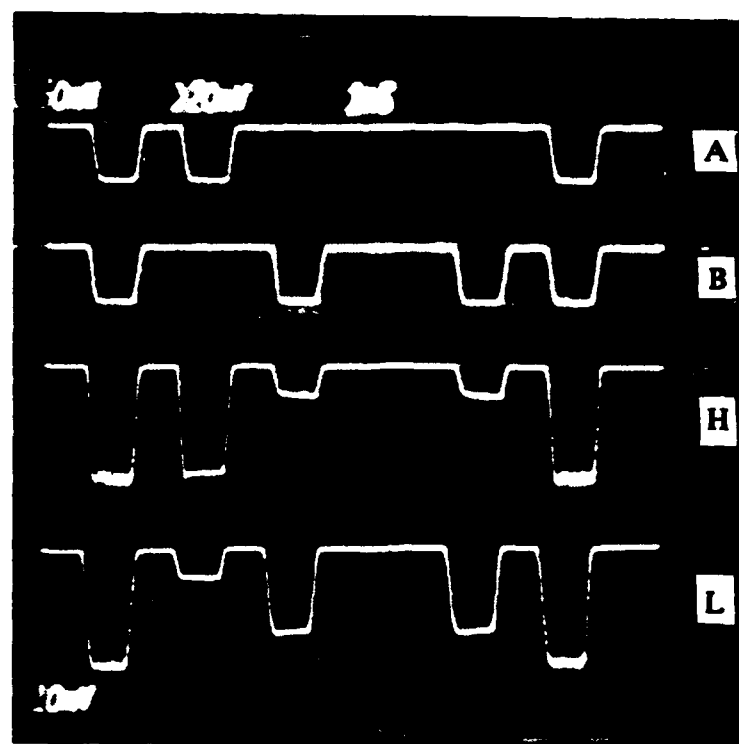


Figure 15.11 The high output and the low output of the system with $A > B$. The input power is 14.5 mW, and the output power is 6.5 mW.

opposite direction will have its polarization rotated in the same direction. Let θ be 45° and let the fast axis of the half-wave plate be 67.5° with respect to the incident plane of vibration of E. The polarization direction of E propagating to the right will be rotated 90° , while that of L which propagates to the left, will be unchanged. A Faraday rotator, as shown in the experimental layout, was not available during the experiment. A half-wave plate was used to rotate the plane of the vibration of E by 90° and to make E go to IF_4 . By detuning the half-wave plate slightly, a small transmission of L is detected.

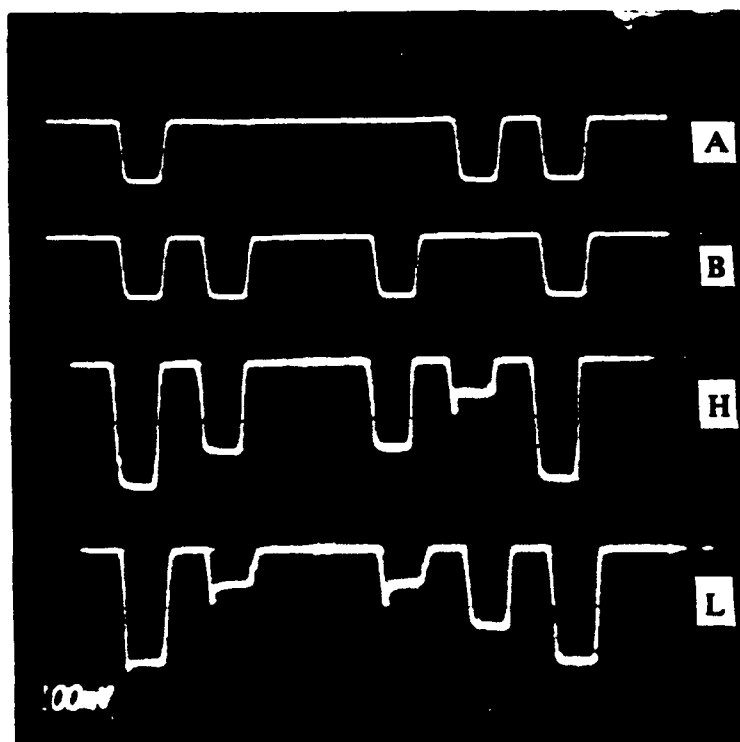


Figure 15.12 The high output and the low output of the system with $B > A$. The input power is 14.5 mW, and the output power is 6.5 mW.

The power of the exchange control signal E in our experiment is small compared to those of numbers A and B. So the exchange operation in our case has to be active. That is, when the exchanger is set to exchange status, it must be switched on for each following bit. The operation speed, then, is limited by this switching speed. However, if E could be twice as large as A and B, the exchange could operate as follows. While $E = 0$, IF_4 cannot be switched on, and always maintains a high reflectivity; the input signals are reflected back, and the exchanger

operates like a mirror. While $E = 1$, IF_4 switches on and maintains a high transmissivity, so that the input signals are transmitted to the opposite sides and exchanged. The advantage is that after the exchanger is set, data can have an extremely high-speed transmission, as routing is linear after the exchange decision has been made.

The polarization isolation is the key to the compare-and-exchange realization. It not only reduces the energy losses in combining signals but also reduces the influences of crosstalk and feedback. The effect of the unused transmissions and reflections has been reduced to a minimum through polarization filtering. Transmissions of A and B through IF_1 and the transmission and reflection of E from IF_4 propagate back toward the source of A and B. The reflection of R_2 from IF_3 reflects directly back. Half of the high transmission of R_1 goes to IF_1 . In this case, however, a decision has been made, and IF_1 is no longer used until the next operation begins. The transmission of $\bar{A}B$ through IF_3 can propagate to IF_2 , which can only be high when an exchange decision is made. Half of the transmission of $\bar{A}B$ from IF_2 can propagate then to IF_3 . This occurrence can be eliminated with an additional Faraday rotator which also prevents \bar{R}_1 from feeding back to IF_1 .

The use of the on-axis normal incidence extends the system to operation on arrays, so that two-dimensional inputs could be compared and exchanged at the same time in parallel.

Using two-dimensional arrays of bistable devices and folded perfect-shuffle interconnections [15.18], optical sorting networks may be feasible for large numbers of channels. First, however, system engineering issues must be addressed. These issues include cascadability, uniformity, crosstalk, reliability, and heat dissipation.

The ZnS interference filters have relatively slow switching times (milliseconds) because they are based on thermal nonlinearities, making real-system applications unlikely. On the other hand, much faster compare-and-exchange modules based on GaAs Fabry-Perot etalons [15.19] may increase the throughput of the sorting networks. Optical latching circuits using GaAs bistable devices have been demonstrated recently [15.20]. GaAs embodiments of the compare-and-exchange designs demonstrated here appear ideal for packet-switching telecommunication networks because GaAs etalons are diode-laser compatible [13.21] and allow rapid reconfiguration of very-high-speed data channels.

In summary, all-optical compare and exchange has been demonstrated using bistable optical devices. The ZnS interference filters used required a speed of 3 ms per bit and a total four-filter power of about 100 mW. The experimental setup can extend to operation on arrays and to other bistable optical devices. The switching

times might be reduced to picoseconds using GaAs etalons, making the system more competitive with alternative approaches.

REFERENCES

- 1.1 H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic Press, New York, 1985).
- 1.2 H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, in *Coherence in Spectroscopy and Modern Physics*, F. T. Arecchi, R. Bonifacio, and M. O. Scully, eds. (Plenum, New York, 1978), p. 111; H. M. Gibbs and N. Peyghambarian, "Optical and hybrid computing," *Proc. SPIE* 634, 142 (1986).
- 1.3 H. M. Gibbs, S. S. Tarng, J. L. Jewell, D. A. Weinberger, K. Tai, A. C. Gossard, S. L. McCall, A. Passner, and W. Wiegmann, *Appl. Phys. Lett.* 41, 221 (1982).
- 1.4 Details were published later in, H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, *Phys. Rev. Lett.* 36, 1135 (1976).
- 1.5 R. Bonifacio and L. A. Lugiato, *Opt. Comm.* 19, 172 (1976).
- 1.6 A. T. Rosenberger, L. A. Orozco, and H. J. Kimble, *Phys. Rev. A* 28, 2569 (1983).
- 1.7 P. W. Smith and E. H. Turner, *Appl. Phys. Lett.* 30, 280 (1977).
- 1.8 E. Garmire, J. H. Margurjer, and S. D. Allen, *Appl. Phys. Lett.* 32, 320 (1978).
- 1.9 D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, *Appl. Phys. Lett.* 45, 13 (1984).
- 1.10 D. A. B. Miller, J. E. Henry, A. C. Gossard, and J. H. English, *Appl. Phys. Lett.* 49, 821 (1986).
- 1.11 J. L. Jewell, M. C. Rushford, and H. M. Gibbs, *Appl. Phys. Lett.* 44, 172 (1984).
- 1.12 J. L. Jewell, Y. H. Lee, M. Warren, H. M. Gibbs, N. Peyghambarian, A. C. Gossard, and W. Wiegmann, *Appl. Phys. Lett.* 46, 918 (1985).
- 1.13 Y. H. Lee, H. M. Gibbs, J. L. Jewell, J. F. Duffy, T. Venkatesan, A. C. Gossard, W. Wiegmann, and J. H. English, *Appl. Phys. Lett.* 49, 486 (1986).
- 1.14 S. D. Smith, A. C. Walker, B. S. Wherrett, F. A. P. Tooley, N. Craft, J. G. H. Mathew, M. R. Taghizadeh, I. Redmond, and R. H. Campbell, *Opt. Eng.* 26, 045 (1987).
- 1.15 L. Wang, H. M. Chou, H. M. Gibbs, G. C. Giglioli, G. Khitrova, H-M. Kulcke, R. Jin, H. A. Macleod, N. Peyghambarian, R. W. Sprague, and M. T. Tsao, *Proc. SPIE O-E Lase 1987*, and *Opt. Eng.* 26, 041 (1987).

- 1.16 L. Wang, V. Esch, R. Feinleib, R. Jin, H. M. Chou, R. W. Sprague, H. A. Macleod, G. Khitrova, H. M. Gibbs, K. Wagner, and D. Psaltis, *Appl. Opt.* 27, 1715 (1988).
- 1.17 L. Zhang, R. Jin, C. W. Stirk, G. Khitrova, R. A. Athale, H. M. Gibbs, H. M. Chou, R. W. Sprague, and H. A. Macleod, *IEEE J. Sel. Areas Commun.* 6, 1273 (1988).
- 2.1 Based in part on the 1988 Ph.D. dissertation of Lon Wang.
- 2.2 M. Sargent III, M. O. Scully, W. E. Lamb, Jr., *Laser Physics* (Addison-Wesley, New York, 1974).
- 2.3 H. M. Gibbs, *Optical Bistability: Controlling Light With Light* (Academic Press, New York, 1985).
- 2.4 J. H. Marburger, and F. S. Felber, *Phys. Rev. A* 17, 335 (1978).
- 2.5 D. A. B. Miller, *Quant. Electron.* 17, 306 (1981).
- 2.6 F. V. Karpushko and G. V. J. Sinitsyn, *Appl. Spectroscopy (USSR)* 29, 1328 (1978).
- 2.7 F. V. Karpushko and G. V. Sinitsyn, *J. Quant. Electron.* 9, 520 (1979).
- 2.8 F. V. Karpushko and G. V. Sinitsyn, *Appl. Phys. B* 28, 137 (1982).
- 2.9 G. R. Olbright, N. Peyghambarian, H. M. Gibbs, H. A. Macleod, and F. Van Milligen, *Appl. Phys. Lett.* 45, 1031 (1984).
- 2.10 A. G. Dirk and H. J. Leamy, *Thin Solid Films* 47, 219 (1977).
- 2.11 H. A. Macleod, *Thin Film Optical Filters, 2nd Edition* (Macmillan, New York, 1986).
- 2.12 E. D. Palik, ed., *Handbook of Optical Constants of Solids* (Academic Press, New York, 1985).
- 2.13 M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1980).
- 2.14 C. R. Pidgeon and S. D. Smith, *J. Opt. Soc. Am.* 54, 1459 (1964).
- 2.15 B. S. Wherrett, *IEEE J. Quant. Electron.* 20, 646 (1984).
- 2.16 B. S. Wherrett, D. Hutchings, and D. Russell, *J. Opt. Soc. Am. B* 3, 351 (1986).
- 2.17 Y. Lee, Ph.D. dissertation (University of Arizona, 1986).
- 2.18 R. J. Harris, G. T. Johnston, G. A. Kepple, P. C. Krok, and H. Mukai, *Appl. Opt.* 16, 436 (1977).
- 2.19 O. Sahlen, *Opt. Commun.* 59, 238 (1986).

- 2.20 E. Abraham, A. K. Kar, M. R. Suttie, R. M. Harris, A. C. Walker, and S. D. Smith, *J. Appl. Phys.* 64, 3393 (1988).
- 2.21 A. C. Walker, *Opt. Commun.* 59, 145 (1986).
- 2.22 Y. T. Chow, B. S. Wherrett, E. V. Stryland, B. T. McGuckin, D. Hutchings, J. G. Mathew, A. Miller, K. and Lewis, *J. Opt. Soc. Am. B* 3, 1535 (1986).
- 3.1 H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985).
- 3.2 N. Peyghambarian and H. M. Gibbs, *J. Opt. Soc. Am. B* 2, 1215 (1985).
- 3.3 For a review see: H. Haug and S. Schmitt-Rink, *Prog. Quant. Electron.* 9, 3 (1984); and "Optical Nonlinearities and Instabilities in Semiconductors," H. Haug, ed. (Academic Press, New York, 1988).
- 3.4 D. S. Chemla and D. A. B. Miller, *J. Opt. Soc. Am. B* 2, 1155 (1985); D. S. Chemla, D. A. B. Miller, P. W. Smith, A. C. Gossard, and W. Wiegmann, *IEEE J. Quantum Electron.* QE-20, 265 (1984).
- 3.5 Y. H. Lee, A. Chavez-Pirson, S. W. Koch, H. M. Gibbs, S. H. Park, J. F. Morhange, A. D. Jeffery, N. Peyghambarian, L. Banyai, A. C. Gossard, and W. Wiegmann, *Phys. Rev. Lett.* 57, 2446 (1986).
- 3.6 H. C. Lee, A. Hariz, P. D. Dapkus, A. Kost, M. Kawase, and E. Garmire, *Appl. Phys. Lett.* 50, 1182 (1987).
- 3.7 J. F. Morhange, S. H. Park, A. D. Jeffery, N. Peyghambarian, H. M. Gibbs, Y. H. Lee, A. Chavez-Pirson, S. W. Koch, A. C. Gossard, J. H. English, M. Maselink, and H. Morkoc, Technical Digest of OSA 1986 Annual Meeting, MT7, Seattle, Washington.
- 3.8 S. Schmitt-Rink, D. S. Chemla, and D. A. B. Miller, *Phys. Rev. B* 32, 6601 (1985).
- 3.9 S. H. Park, J. F. Marhange, A. D. Jeffery, R. A. Morgan, A. Chavez-Pirson, H. M. Gibbs, S. W. Koch, N. Peyghambarian, M. Derstine, A. C. Gossard, J. H. English, and W. Wiegmann, *Appl. Phys. Lett.* 52, 1201 (1988).
- 4.1 Based in part on the 1987 M.S. Thesis of Hsiang-Ming Chou (University of Arizona).
- 4.2 Based in part on Chapter 6 of the 1989 Ph.D. dissertation of Arturo Chavez-Pirson (University of Arizona).
- 4.3 J. L. Jewell, A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, *Appl. Phys. Lett.* 51, 94 (1987).
- 4.4 O. Sahlen, U. Olin, E. Masseboeuf, G. Landgren, and M. Rask, *Appl. Phys. Lett.* 22, 1559 (1987).
- 4.5 J. P. van der Ziel and M. Ilegems, *Appl. Opt.* 14, 2627 (1975).

- 4.6 R. L. Thornton, R. D. Burnham, and W. Streifer, Appl. Phys. Lett. 45, 1028 (1984).
- 4.7 P. L. Gourley and T. J. Frummond, Appl. Phys. Lett. 49, 489 (1986).
- 4.8 R. Kuszelewicz, J. L. Oudar, J. C. Michel, and R. Azoulay, Appl. Phys. Lett. 55, 2138 (1988).
- 4.9 M. Warren, "Fabrication, Experimental Investigation and Computer Modeling of Gallium Arsenide Nonlinear Optical Devices," Ph.D. dissertation (University of Arizona, 1988).
- 4.10 Y. H. Lee, H. M. Gibbs, J. L. Jewell, J. F. Duffy, T. N. C. Venkatesan, A. C. Gossard, W. Wiegmann, and J. H. English, Appl. Phys. Lett. 49, 486 (1986).
- 4.11 J. L. Jewell, "Fabrication, Investigation, and Optimization of GaAs Optical Bistable Devices and Logic Gates," Ph.D. dissertation (University of Arizona, 1984).
- 4.12 U. Olin and O. Sahlen, Conference on Lasers and Electro-Optics (CLEO), Baltimore, Maryland, Paper TuJ33 (1989).
- 5.1 H. M. Gibbs, S. L. McCall, T. N. C. Venkatesan, A. C. Gossard, A. Passner, and W. Wiegmann, Appl. Phys. Lett. 35, 451 (1979).
- 5.2 H. M. Gibbs, S. S. Tarng, J. L. Jewell, D. A. Weinberger, K. Tai, A. C. Gossard, S. L. McCall, A. Passner, and W. Wiegmann, Appl. Phys. Lett. 41, 221 (1982).
- 5.3 H. M. Gibbs, *Optical Bistability: Controlling Light with Light*, Academic Press, New York (1985). Also N. Peyghambarian and H. M. Gibbs, Opt. Eng. 24(1), 68 (1985), S. W. Koch, N. Peyghambarian, and H. M. Gibbs, Appl. Phys. Rev. (January 1988).
- 5.4 S. S. Tarng, J. Jewell, N. Peyghambarian, A. C. Gossard, T. Venkatesan, and W. Wiegmann, Appl. Phys. Lett. 44, 360 (1984).
- 5.5 S. S. Tarng, K. Tai, J. L. Jewell, H. M. Gibbs, A. C. Gossard, S. L. McCall, A. Passner, T. N. C. Venkatesan, and W. Wiegmann, Appl. Phys. Lett. 40, 205 (1982).
- 5.6 J. L. Jewell, Y. H. Lee, J. F. Duffy, A. C. Gossard, W. Wiegmann, and J. H. English, in *Optical Bistability III*, H. M. Gibbs, P. Mandel, N. Peyghambarian, and S. D. Smith, eds. (Springer-Verlag, Berlin, 1985), pp. 32-34.
- 5.7 M. Warren, Y. H. Lee, G. R. Olbright, B. P. McGinnis, H. M. Gibbs, N. Peyghambarian, T. Venkatesan, B. Wilkens, J. Smith, and A. Yariv, in *Optical Bistability III*, H. M. Gibbs, P. Mandel, N. Peyghambarian, and S. D. Smith, eds. (Springer-Verlag, Berlin, 1985), pp. 39-41.
- 5.8 J. L. Jewell, A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, PDP1-1, Topical Meeting on Photonic Switching, Incline Village, Nevada (1987).

- 5.9 R. Jin, L. Wang, R. W. Sprague, H. M. Gibbs, G. C. Giglioli, H. Kulcke, H. A. Macleod, N. Peyghambarian, G. R. Olbright, and M. Warren, in *Optical Bistability III*, H. M. Gibbs, P. Mandel, N. Peyghambarian, and S. D. Smith, eds. (Springer-Verlag, Berlin, 1985).
- 5.10 M. T. Tsao, L. Wang, R. Jin, R. W. Sprague, G. Giglioli, H. M. Kulcke, Y. D. Li, H. M. Chou, H. M. Gibbs, and N. Peyghambarian, *Opt. Eng.* 26(1), 41 (1987).
- 5.11 F. A. P. Tooley, S. D. Smith, and C. T. Seaton, *Appl. Phys. Lett.* 43, 807 (1983).
- 5.12 D. A. B. Miller, S. D. Smith, and A. Johnston, *Appl. Phys. Lett.* 35, 658 (1979).
- 5.13 B. K. Jenkins, A. A. Sawchuk, T. C. Strand, R. Forchheimer, and B. H. Soffer, *Appl. Opt.* 23, 3455 (1984).
- 5.14 S. D. Smith, I. Janossy, H. A. MacKenzie, J. G. H. Mathew, J. J. E. Reid, M. R. Taghizadeh, F. A. B. Tooley, and A. C. Walker, *Opt. Eng.* 24(4), 569 (1985).
- 5.15 L. Wang, V. Esch, R. Feinleib, L. Zhang, H. M. Chou, R. W. Sprague, H. A. Macleod, G. Khitrova, H. M. Gibbs, K. Wagner, and D. Psaltis, *Appl. Opt.* 27, 1715 (1988).
- 5.16 A. Huang, *Proc. IEEE 10th International Optical Computing Conference*, pp. 13-17, publication No. 83CH1880-4 (1983).
- 5.17 L. Zhang, R. Jin, C. W. Stirk, G. Khitrova, R. A. Athale, H. M. Gibbs, H. M. Chou, R. W. Sprague, and H. A. Macleod, *IEEE J. Sel. Areas in Commun.* 6, 1273 (1988).
- 5.18 J. Hegarty and K. A. Jackson, *Appl. Phys. Lett.* 45(12), 1314 (1984).
- 6.1 T. Bischofberger and Y. R. Shen, *Phys. Rev. A* 19, 1169 (1979).
- 6.2 L. Banyai and S. W. Koch, *Z. Physik. B* 63, 283 (1986).
- 6.3 Y. H. Lee, A. Chavez-Pirson, S. W. Koch, H. M. Gibbs, S. H. Park, J. Morhange, N. Peyghambarian, L. Banyai, A. C. Gossard, and W. Wiegmann, *Phys. Rev. Lett.* 57, 2446 (1987).
- 6.4 Y. H. Lee, H. M. Gibbs, J. L. Jewell, J. F. Duffy, T. Venkatesan, A. C. Gossard, W. Wiegmann and J. H. English, *Appl. Phys. Lett.* 49, 486 (1986).
- 6.5 Y. H. Lee, M. Warren, G. R. Olbright, H. M. Gibbs, N. Peyghambarian, T. Venkatesan, J. S. Smith and A. Yariv, *Appl. Phys. Lett.* 48, 754 (1986).
- 6.6 J. L. Jewell, A. Scherer, S. L. McCall, A. C. Gossard, and J. H. English, *PDP1-1, Topical Meeting on Photonic Switching, Incline Village, Nevada* (1987).
- 6.7 N. Peyghambarian and S. W. Koch, *Rev. Phys. Appl.* 22, 1711 (1987).
- 7.1 R. Jin, C. Hanson, M. Warren, D. Richardson, H. M. Gibbs, N. Peyghambarian, G. Khitrova, and S. W. Koch, *Appl. Phys. B* 45, 1 (1988).

- 7.2 Ch. Golaire, Paul Mandel, and P. Laws, *Opt. Commun.* 57, 297 (1986).
- 7.3 R. Cush and I. Bennion, TuC5-1, Topical Meeting on Optical Computing, Incline Village, Nevada (1987).
- 7.4 F. A. P. Tooley, *Appl. Opt.* 26, 1741 (1987).
- 7.5 L. Banyai and S. W. Koch, *Z. Physik B* 63, 283 (1986).
- 7.6 See, e.g., the article by S. W. Koch in *Nonlinear Properties of Semiconductors*, H. Haug, ed. (Academic, New York, 1988).
- 7.7 A. Yariv, *Optical Electronics*, (CBS College Publishing, New York, 1985), pp. 88-89.
- 7.8 H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985).
- 7.9 M. E. Warren, S. W. Koch, and H. M. Gibbs, Invited paper for special issue of IEEE Computer Society publication "Computer," T. E. Bachman and E. A. Parrish, eds. (1987).
- 7.10 J. Gaskill, *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978).
- 8.1 J. W. Goodman, *Optica Acta* 32, 1489 (1985).
- 8.2 J. R. Leger, G. L. Swanson, W. B. Veldkamp, *Appl. Opt.* 26, 4391 (1987).
- 9.1 S. H. Lin, S. Y. Wang, S. A. Newton, Y. M. Houng, *Electron. Lett.* 21, 597 (1985).
- 9.2 S. R. Friberg, L. Zhang, C. F. Li, paper TuK4, OSA Annual Meeting, Santa Clara, California, Oct. 1988.
- 9.3 T. Venkatesan, P. J. Lemaire, B. Wilkens, L. Soto, A. G. Gossard, W. Wiegmann, J. L. Jewell, H. M. Gibbs, and S. S. Tarn, *Opt. Lett.* 9, 297 (1984).
- 9.4 L. Zhang, R. Jin, C. W. Stirk, G. Khitrova, R. A. Athale, H. M. Gibbs, H. M. Chou, R. W. Sprague, and H. A. Macleod, *IEEE Journal on Select. Areas in Commun.* 6, 1273 (1988).
- 9.5 R. Jin, C. Hanson, M. Warren, D. Richardson, H. M. Gibbs, N. Peyghambarian, G. Khitrova, and S. W. Koch, *Appl. Phys. B* 46, 61 (1988).
- 9.6 Y. H. Lee, A. Chavez-Pirson, B. K. Rhee, H. M. Gibbs, A. C. Gossard, and W. Wiegmann, *Appl. Phys. Lett.* 49, 1505 (1986).
- 9.7 K. Tai, J. L. Jewell, W. T. Tsang, H. Temkin, M. Panish, and Y. Twu, *Appl. Phys. Lett.* 50, 795 (1987).
- 10.1 Based on Chapter 6 of the 1988 Ph.D. Dissertation of Lon Wang (University of Arizona).

- 10.2 D. Psaltis and N. Farhat, *Opt. Lett.* 10, 98 (1985).
- 10.3 D. Psaltis, J. Yu, X. Gu, and H. Lee, *Technical Digest of Topical Meeting on Optical Computing* 11, 129 (1987).
- 10.4 A. W. Lohmann, *Appl. Opt.* 25, 1530 (1986).
- 10.5 M. J. Murdocca and N. Streibl, *Technical Digest of Topical Meeting on Optical Computing* 11, 9 (1987).
- 10.6 W. H. Knox, R. L. Fork, M. C. Downer, R. H. Stolen, C. V. Shank, and J. S. Valdamanis, *Appl. Phys. Lett.* 46, 1120 (1985).
- 10.7 M. R. Pinnel and W. H. Knausenberger, *AT&T Tech. J.* 60, 45 (1987).
- 10.8 A. J. Blodgett, Jr., *Sci. Amer.* 249, 86 (1983).
- 10.9 C. W. Hoover, Jr., W. L. Harrod, M. I. Cohen, *AT&T Tech. J.* 66, 1 (1987).
- 10.10 J. R. Brews, *IEEE Electron Device Letters* 1, 2 (1980).
- 10.11 A. Reisman, *Proc. IEEE* 71, 555 (1983).
- 10.12 K. C. Saraswat and F. Mohammadi, *IEEE J. Solid-State Circuits* SC-17, 275 (1982).
- 10.13 S. Seki and H. Hasegawa *IEEE Trans. Microwave Theory Tech.* 32, 1715 (1984).
- 10.14 A. M. Wilson, *Thin Solid Films* 83, 145 (1981).
- 10.15 A. J. Rainal, *Bell Syst. Tech. J.* 995 (May-June 1979).
- 10.16 Antinone, R., *IEEE Journal of Solid State Circuits* SC-18, 434 (1987).
- 10.17 J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale, *Proc. IEEE* 72, 850 (1984).
- 10.18 R. K. Kostuk, Ph.D. dissertation (Stanford Univ. 1986).
- 10.19 S. M. Arnold, *Opt. Eng.* 24, 803 (1985).
- 10.20 K. Winick, *J. Opt. Soc. Am.* 72, 143 (1982).
- 10.21 H. Bartelt and S. K. Case, *Appl. Opt.* 21, 2886 (1982).
- 10.22 S. Shammash and P. A. Goud, *Opt. Lett.* 6, 222 (1981).
- 10.23 K. Yokomori, *Appl. Opt.* 23, 2303 (1984).
- 10.24 T. K. Gaylord and G. Moharam, *Proc. IEEE* 73, 894 (1985).

- 10.26 H. Dammann and E. Klotz, *Optica Acta* 24, 505 (1977).
- 10.27 W.-H. Lee, *Appl. Opt.* 18, 2152 (1979).
- 10.28 W.-H. Lee, *Progress in Optics, Volume 16*, E. Wolf, ed. (North Holland, Amsterdam, 1979).
- 10.29 U. Killat, G. Rabe, and W. Rave, *Fiber and Integrated Optics*, 2, 159 (1982).
- 10.30 J. R. Leger, G. J. Swanson, and W. B. Veldkamp, *Appl. Opt.* 26, 4391 (1987).
- 10.31 W.-H. Lee, *Opt. Commun.* 36, 469 (1981).
- 10.32 E. Hecht and A. Zajac, *Optics* (Addison-Wesley, New York, 1974).
- 10.33 J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
- 10.34 T. R. Cuthbert, Jr., *Optimization Using Personal Computers* (John Wiley & Sons, New York, 1986).
- 10.35 H. Kogelnik, *Bell Syst. Tech. J.* 48, 2909 (1969).
- 10.36 T. Jaaskelainen and T. Hytonen, *Opt. Commun.* 64, 19 (1987).
- 10.37 B. J. Chang, and C. D. Leonard, *Appl. Opt.* 18, 2407 (1979).
- 10.38 P. Hariharan, *Optical Holography* (Cambridge University Press, Cambridge, 1984).
- 10.39 R. R. A. Syms and L. Solymar, *Appl. Opt.* 22, 1479 (1983).
- 10.40 S. K. Case and R. Alferness, *Appl. Phys.* 10, 41 (1976).
- 10.41 D. H. Close and A. Graube, Technical Report AFML-TR-73-267 (1973).
- 10.42 J. L. Horner and J. E. Ludman, *Appl. Opt.* 20, 1845 (1981).
- 10.43 J. T. LaMacchia and C. J. Vincelette, *Appl. Opt.* 7, 1857 (1968).
- 10.44 S. K. Case, Ph.D. dissertation (University of Michigan, 1976).
- 10.45 R. Kowarschik, *Opt. Quant. Elect.* 10, 171 (1978).
- 10.46 G. Gigioli, Ph. D. dissertation (University of Arizona, 1988).
- 11.1 L. Wang, W.-Z. Geng, and R. A. Kostuk, *Proc. SPIE*, to be published.
- 12.1 E. Abraham, A. K. Kar, M. R. Suttie, R. M. Harris, A. C. Walker, and S. D. Smith, *J. Appl. Phys.* 64, 3393 (1988).
- 12.2 R. Jin, C. L. Chuang, H. M. Gibbs, S. W. Koch, J. N. Polky, and G. A. Pubanz, *Appl. Phys. Lett.* 53, 1791 (1988).

- 12.3 For Bell Labs work on microresonators see J. L. Jewell, S. L. McCall, Y. H. Lee, A. Scherer, A. C. Gossard, and J. H. English, *J. de Physique, Colloque* C2 49, C2-39 (1988).
- 12.4 1988 MS thesis of Richard Eric Feinleib, (University of Arizona) "Adaptive Optical Learning Network with a Photorefractive Crystal," describes incoherent and coherent erasure of gratings in LiNbO_3 , BaTiO_3 and SBN.
- 12.5 G. Khitrova, D. Rouede, N. V. Kukhtarev, and H. M. Gibbs, *Phys. Rev. Lett.* 62, 1110 (1989).
- 13.1 Based in part on Chapter 4 of the Ph.D. dissertation of Lon Wang (University of Arizona, 1988).
- 13.2 A. Huang, in *IEEE 10th International Optical Computing Conference Technical Digest*, 13 (1983).
- 13.3 K. H. Brenner, *Appl. Opt.* 25, 3060 (1986).
- 14.1 Based on Chapter 5 of the Ph.D. dissertation of Lon Wang (University of Arizona, 1988).
- 14.2 B. H. Soffer, G. J. Dunning, Y. Owechko, and E. Marom, *Opt. Lett.* 11, 118 (1986).
- 14.3 J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* 79, 2554 (1982).
- 14.4 T. Kohonen, *Self-Organization and Associative Memory* (Springer-Verlag, New York, 1984).
- 14.5 D. Psaltis, and N. Farhat, *Opt. Lett.* 10, 98 (1985).
- 14.6 Y. S. Abu-Mostafa and D. Psaltis, *Sci. Amer.* 256, 88 (1987).
- 14.7 E. G. Paek and D. Psaltis, *Opt. Eng.* 26, 428 (1987).
- 14.8 D. Anderson and M. C. Erie, *Opt. Eng.* 26, 434 (1987).
- 14.9 Y. Owechko, J. Dunning, E. Marom, and B. H. Soffer, *Appl. Opt.* 26, 1900 (1987).
- 14.10 G. R. Mangasaryan, B. E. Khaikin, and V. S. Khitrova, *Sov. J. Quantum Electron.* 6, 318 (1976).
- 14.11 T. G. Georgekutty and H.-K. Liu, *Appl. Opt.* 26, 371 (1987).
- 14.12 L. H. Lin, *J. Opt. Soc. Am.* 61, 203 (1971).
- 14.13 J. W. Goodman, *Introduction to Fourier Optics*, (McGraw-Hill, New York, 1968).
- 14.14 J. Hong and D. Psaltis, *Opt. Lett.* 11, 812 (1986).

- 15.1 D. E. Knuth, *The Art of Computer Programming: Sorting and Searching, Vol. 3* (Addison-Wesley, Reading, Massachusetts, 1973).
- 15.2 S. G. Akl, *Parallel Sorting Algorithms* (Academic Press, Orlando, Florida, 1985).
- 15.3 A. Borodin and J. E. Hopcroft, *J. Comput. Syst. Sci.* 30, 130 (1985).
- 15.4 T. Leighton, in *Proc. 16th Annual ACM Symp. Theory Comput.*, 1984, p. 71.
- 15.5 K. E. Batcher, in *Proc. 1968 Spring Joint Comput. Conf.*, 32. (AFIPS Press, Reston, Virginia), pp. 307-314.
- 15.6 H. S. Stone, *IEEE Trans. Comput.* C-20, 153 (1971).
- 15.7 A. W. Lohmann, *Appl. Opt.* 25, 1543 (1986).
- 15.8 J. E. Midwinter, *IEE Proc.* 132, Pt. J, 371 (1985).
- 15.9 C. W. Stirk, R. A. Athale, and M. W. Haney, *Appl. Opt.* 27, 202 (1988).
- 15.10 A. Huang, in *Proc. Int. Conf. Commun.*, Toronto, Ontario, Canada, June 22, 1986.
- 15.11 H. P. Moravec, *IEEE Trans. Comput.* C-28, 795 (1979).
- 15.12 W. I. Shu and A. K. Sood, in *Proc. Vector Parallel Processors Computat. Sci. II* (Oxford, U.K., Aug. 1984).
- 15.13 C. W. Stirk, R. A. Athale, and C. B. Friedlander, *Appl. Opt.*, submitted for publication.
- 15.14 H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic Press, New York, 1985).
- 15.15 H. M. Gibbs and N. Peyghambarian, *Proc. SPIE* 634, 142 (1987).
- 15.16 S. D. Smith, I. Janossy, H. A. MacKenzie, J. G. H. Mathew, J. J. E. Reid, M. R. Taghizadeh, F. A. P. Todey, and A. C. Walker, *Opt. Eng.* 24, 569 (1985).
- 15.17 L. Wang, H. M. Chou, H. M. Gibbs, G. C. Giglioli, G. Khitrova, H.-M. Kulcke, R. Jin, H. A. Macleod, N. Peyghambarian, R. W. Sprague, and M. T. Tsao, *Proc. SPIE* 752 (1987).
- 15.18 C. W. Stirk, R. A. Athale, and C. B. Friedlander, *Proc. SPIE* 754, 175 (1987).
- 15.19 Y. H. Lee, H. M. Gibbs, J. L. Jewell, J. F. Duffy, T. Venkatesan, A. C. Gossard, W. Wiegmann, and J. H. English, *Appl. Phys. Lett.* 49, 486 (1986).
- 15.20 R. Jin, C. Hanson, M. Warren, D. Richardson, H. M. Gibbs, N. Peyghambarian, G. Khitrova, and S. W. Koch, *Appl. Phys. B* 46, 61 (1988).
- 15.21 A. Migus, A. Antonetti, D. Hulin, A. Mysyrowicz, H. M. Gibbs, N. Peyghambarian, and J. L. Jewell, *Appl. Phys. Lett.* 46, 70 (1985).

APPENDIX A. ROOM-TEMPERATURE EXCITONIC OPTICAL NONLINEARITIES OF MOLECULAR BEAM EPITAXIALLY GROWN ZnSe THIN FILMS

A.1 Introduction

From the time optical bistability was first in GaAs and in InSb, a variety of semiconductor materials has been considered for application to nonlinear signal processing and to optical computing. Interference filters of polycrystalline ZnS and ZnSe have been studied extensively, because they are easy to fabricate, and because they can operate at room temperature and in the visible wavelength range [A.1-A.3]. Thermal nonlinearities have been employed in these interference filters, however, resulting in inefficient switching with slow switching speeds [A.2, A.3]. Nonlinear excitonic effects at $T = 77$ K also have been reported in ZnSe/ZnMnSe superlattices [A.4]. We have observed large nonlinearities of electronic origin in thin films of ZnSe grown by molecular beam epitaxy (MBE) at low and room temperatures [A.5, A.6]. The nonlinear index of refraction also has been obtained.

Zinc selenide is a direct gap semiconductor, with a zinc blende crystal structure and a relatively large band-gap energy of $E_g \approx 2.7$ eV at $T = 300$ K, corresponding to the visible blue part of the spectrum. Excitons in ZnSe are relatively strongly bound with a binding energy $E_{Bx} \approx 18$ meV and a Bohr radius ≈ 51 Å [A.7]. For optical bistability applications and for optimization of logic devices, the material must be characterized. One must understand the origin of optical nonlinearities and know the nonlinear refractive index over a range of wavelengths in the vicinity of the band gap. Such measurements have been performed and are described here.

A.2 Experimental Results

The standard pump-and-probe method was used to determine the change in the absorption coefficient as a function of probe frequency for various pump intensities. The broad-band spontaneous luminescence emitted by a cell containing Coumarin-460 was used for the probe beam. The pump beam was generated by a tunable dye laser using Coumarin-440. Both dye laser and dye cell were pumped synchronously by the same nitrogen laser. Both beams were ≈ 3 ns full-width at half-maximum (FWHM), and were collimated and brought onto the sample so as to be coincident in time and space. The spot diameter of the pump beam on the sample was ≈ 100 μ m while the probe beam diameter was ≈ 70 μ m. The wavelength of the pump beam was fixed at 435 nm, generating an electron-hole plasma rather than excitons. The

pump beam was polarized orthogonally with respect to the probe beam which allowed a selective blocking of the pump beam before it entered the spectrometer. Only the transmitted probe beam was collected by the optical multichannel analyzer (OMA). The sample was grown by MBE on (100) GaAs substrates and was 0.55- μm thick. The GaAs substrate was etched away for the purpose of the experiment. Samples for this study were grown by MBE from elemental Zn and Se sources at a growth temperature of 350° C and a Zn/Se beam pressure ratio of 1:3. Further details concerning the growth and characterization of these samples are found in the literature [A.8, A.9].

Figure A.1(a) shows the absorption coefficient of the probe as a function of probe wavelength for various pump intensities at $T = 150$ K. The spectrum labeled 0, which is taken in the absence of the pump beam, exhibits a pronounced exciton peak around 450 nm, and Coulomb-enhanced continuum states at higher energies. In the presence of the pump pulse, one observes a reduction of the exciton oscillator strength and a small broadening on either side of the exciton. At higher pump intensities the exciton gradually disappears. This effect is caused by screening of the Coulomb potential by the pump-injected carriers. The change in the index of refraction, Δn , can be obtained by making a Kramers-Kronig transformation of the measured change in the absorption coefficient, $\Delta\alpha$. Figure A.1(b) shows the nonlinear index changes corresponding to measured absorption spectra. The maximum Δn is approximately -0.025 at about $1/2E_{Bx} \approx 9$ meV below the exciton energy with ≈ 2.5 kW/cm² pump intensity.

Similar measurements of $\Delta\alpha$ and Δn were taken at room temperature, as shown in Figs. A.2(a) and A.2(b). The exciton resonance is broadened at this temperature by phonon interaction. Again, screening of the exciton resonance is observed with some collision broadening as the pump intensity is increased. The maximum Δn at room temperature is -0.017 at about $1E_{Bx} \approx 18$ meV below the exciton energy with ≈ 16 kW/cm² pump intensity.

A.3 Discussion of Results

The experimental data were analyzed using recently-developed plasma theory [A.10]. This theory has been employed successfully to analyze experimental findings for bulk GaAs [A.6] and CdS_xSe_{1-x}-doped glasses [A.11]. It takes into account band filling, screening of the Coulomb interaction, and density-dependent broadening of resonances. The absorption coefficient is written as

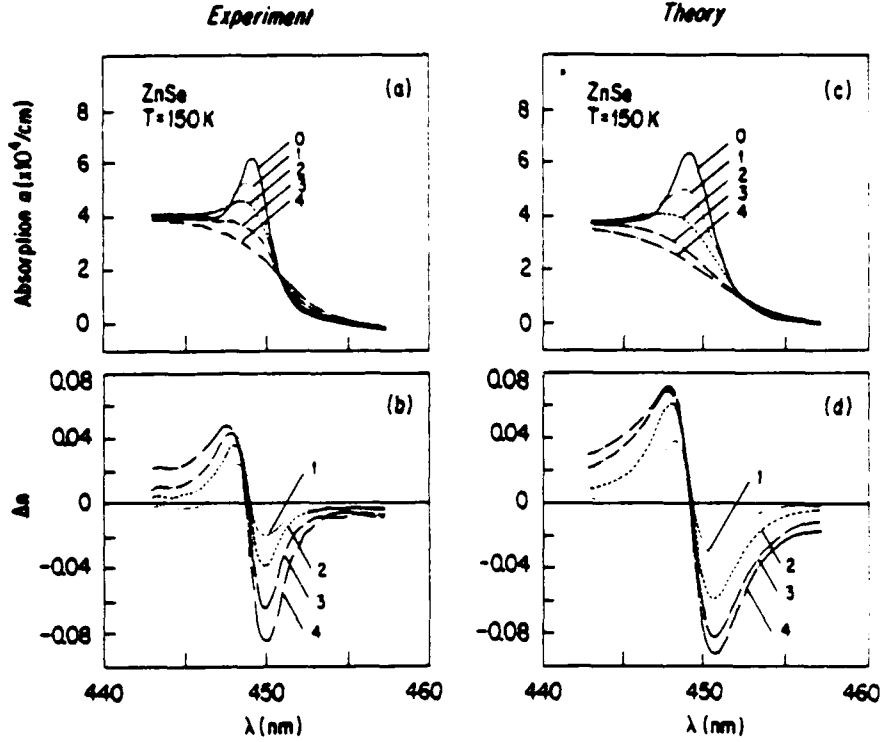


Figure A.1 Experimental and theoretical spectra of nonlinear absorption and nonlinear refractive index for a 0.55- μm ZnSe thin film at $T = 150$ K. (a) Experimental absorption spectra for various pump intensities: (0) linear (no pump), (1) 2.5 kW/cm², (2) 9 kW/cm², (3) 23 kW/cm², (4) 46.5 kW/cm². (b) nonlinear refractive index changes corresponding to the measured absorption spectra of (a). Curves 1-4 are obtained from a Kramers-Kronig transformation of the corresponding experimental data 1-4 in (a). (c) Calculated absorption spectra for different electron-hole pair densities, N : (0) $1 \times 10^{15} \text{ cm}^{-3}$, (1) $5 \times 10^{16} \text{ cm}^{-3}$, (2) $1 \times 10^{17} \text{ cm}^{-3}$, (3) $2 \times 10^{17} \text{ cm}^{-3}$, (4) $3 \times 10^{17} \text{ cm}^{-3}$. (d) Calculated nonlinear refractive index changes, curves 1-4 are obtained from curves 1-4 in (c), respectively.

$$\alpha(\omega) = c \tanh \left[\frac{\hbar\omega - \mu}{2k_B T} \right]$$

$$X \sum_{\lambda} |\phi_{\lambda}(r=0)|^2 \delta_r(\hbar\omega - E_{\lambda} - E_g) \quad . \quad (\text{A.1})$$

where c is a constant proportional to the absolute square of the interband matrix element, $\mu = \mu_e + \mu_h$ is the total chemical potential of the excited electron-hole pair,

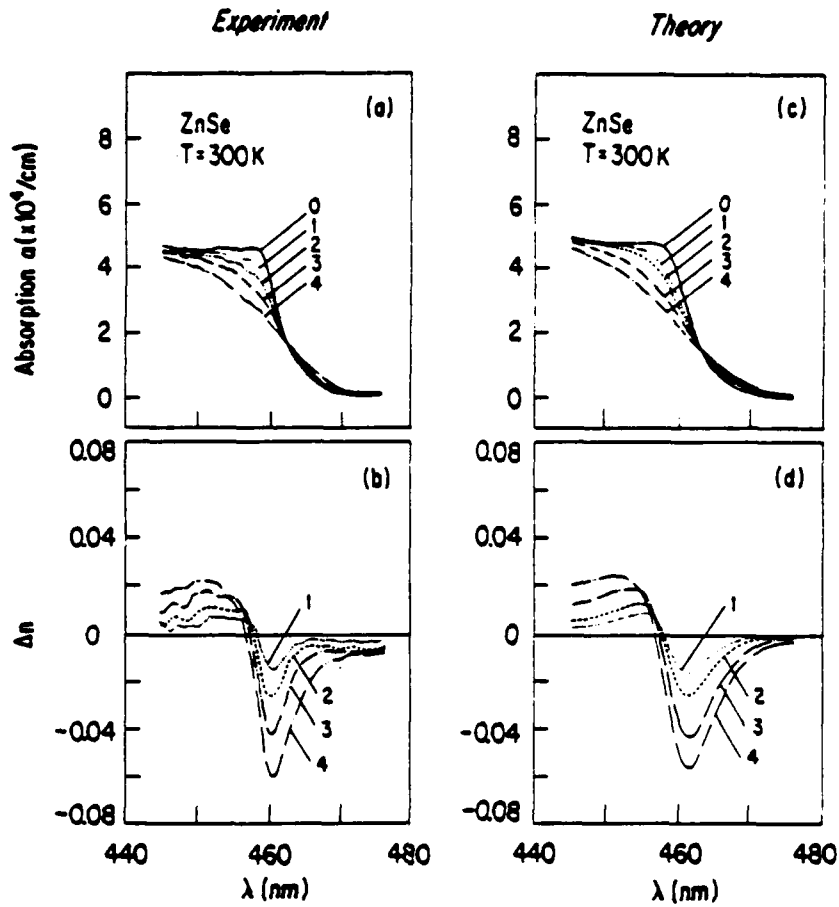


Figure A.2 Experimental and theoretical spectra of the nonlinear absorption and nonlinear refractive index for a 0.55- μm ZnSe thin film at room temperature. (a) experimental absorption spectra for various pump intensities: (0) no pump, (1) 16 kW/cm^2 , (2) 33 kW/cm^2 , (3) 76 kW/cm^2 , and (4) 157 kW/cm^2 . (b) Nonlinear refractive index changes corresponding to the measured absorption spectra; curves 1-4 are obtained from curves 1-4 in (a), respectively. (c) Calculated spectra for different electron-hole pair densities, N : (0) 10^{15} cm^{-3} , (1) $9 \times 10^{16} \text{ cm}^{-3}$, (2) $1.5 \times 10^{17} \text{ cm}^{-3}$, (3) $3 \times 10^{17} \text{ cm}^{-3}$, and (4) $5 \times 10^{17} \text{ cm}^{-3}$. (d) Calculated nonlinear refractive index changes, curves 1-4 are obtained from curves 1-4 in (c), respectively.

T is temperature, and $\phi_\lambda(r)$ is the wave function for the relative motion of an electron-hole pair (either bound, i.e., exciton, or unbound). $\phi_\lambda(r)$ is the solution of the Wannier equation

$$\left[-\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right] \phi_\lambda(r) = E_\lambda \phi_\lambda(r) . \quad (\text{A.2})$$

where $V(r)$ is the screened Coulomb potential and E_λ is the energy eigenvalue. $E_g = E_g^0 - \delta E_g$ is the renormalized band-gap energy, with δE_g being the shift of the

band gap to lower energies (band-gap renormalization due to intraband exchange and correlation effects) and E_g^0 the unrenormalized gap energy. The broadened δ function $\delta_\Gamma(x)$ in Eq. (A.1) is a consequence of the conservation of energy. It is taken as a broadened line shape function to simulate the Urbach tail of the excitons:

$$\delta_\Gamma(x) = \frac{E_R}{\pi\Gamma \cosh(xE_R/\Gamma)} \quad (A.3)$$

Here $\Gamma = \Gamma_0 + \Gamma_1 N$ is the width of the line shape function. It has density-dependent and density-independent terms, Γ_1 and Γ_0 , respectively, with N being the injected carrier density. N is related to the measured light intensity I , via the simple rate equation (in steady state)

$$\begin{aligned} \frac{dN}{dt} &= \frac{\alpha(\omega, N)I}{\hbar\omega} - \frac{N}{\tau} \\ N &= \frac{\alpha(\omega, N)I\tau}{\hbar\omega} \end{aligned}$$

The hyperbolic tangent term in Eq. (A.1) is a consequence of the band filling, which in turn increases transmission at lower energies. The increase in transmission is attributable to the Pauli exclusion principle, which blocks interband transitions for filled states. Screening, which is incorporated in the potential energy $V(r)$, reduces the exciton oscillator strength, the exciton binding energy, and the Coulomb enhancement of the continuum states.

Figure A.1(c) shows the calculated absorption spectra for various electron-hole pair densities, N , at $T = 150$ K. The spectrum labeled (0), which is calculated for $N = 1 \times 10^{15} \text{ cm}^{-3}$, represents the linear absorption, demonstrating that lower carrier densities do not change the shape of the spectrum. The value of c in Eq. (A.1) is adjusted to reproduce the measured continuum absorption coefficient (for a frequency range above the exciton resonance). The value of $\Gamma_0 = 0.4E_R$ also is adjusted to reproduce the correct ratio of exciton peak absorption to continuum absorption. Once these parameters are obtained for the linear spectrum, they are held fixed for the rest of the calculation. The spectrum labeled (1) in Fig. A.1(c) is calculated for $N = 5 \times 10^{16} \text{ cm}^{-3}$. A value of $\Gamma_1 = (2.5 \times 10^{-18} \text{ cm}^{-3})E_R$ was used to reproduce the broadening observed at higher pump intensities. As the density of carriers is increased, the exciton is screened and broadened, as shown in Fig. A.1(c). The calculated nonlinear refractive index changes obtained from the spectra of Fig. A.3(c) are plotted in Fig. A.1(d).

Similar calculations have been performed at room temperature, and the results are plotted in Figs. A.2(c) and A.2(c) for absorption and refractive index changes, respectively. The parameters used are $\Gamma_0 = E_R$ and $\Gamma_1 = (2.5 \times 10^{-18} \text{ cm}^{-3})E_R$. In order to determine the relative importance of the effects of band filling, screening, and broadening in the measured nonlinear index, we may artificially turn off an effect in the calculations and then compare the obtained magnitude and shape of the nonlinear index with the measured quantities [A.12]. Analysis of the spectra at room temperature (Fig. A.2) shows that the main contributions to the nonlinear index are exciton screening and density-dependent broadening. Turning off the band filling term produces spectra that closely resemble the measured spectra, indicating that band filling effects are unimportant at the intensities employed. This result is in contrast to room-temperature optical nonlinearities of bulk GaAs, in which approximately 50% of the nonlinearity is attributed to band filling and 50% to screening [A.12]. This contrast is attributable to the much larger exciton binding energy in ZnSe as compared to that in GaAs. Band-filling effects usually have a dominant contribution after the exciton has been ionized. In ZnSe, the Mott density at room temperature is $\sim 4 \times 10^{17} \text{ cm}^{-3}$ (it is $\cong 2 \times 10^{17} \text{ cm}^{-3}$ for $T = 150 \text{ K}$), so that at the highest intensities used ($N \sim 5 \times 10^{17} \text{ cm}^{-3}$), the exciton is just ionized. Only at higher densities does band filling begin to play a major role.

At lower temperatures, the band filling term becomes even less important, because the Fermi distributions become sharper. At $T = 150 \text{ K}$, we see that the major contribution to nonlinearity is screening of the exciton. Band-gap renormalization at these densities does not play an important role; it becomes most effective after the exciton has been completely ionized, i.e., after the band gap shifts through the former exciton resonance. It must be emphasized that in contrast to most previous results on ZnSe, thermal effects do not play a role in the observed spectra. This is because the pulses are 3 ns long at a 10-Hz repetition rate.

A.4 Measurement of the Response Time of the Optical Nonlinearities in ZnSe

In the second set of experiments we measured the dynamics of the excitonic nonlinearities using time-resolved pump-probe spectroscopy. Free carriers were created by pumping above the ZnSe band edge with a 310-nm pump. The $\cong 100$ -fs pump pulse was obtained by frequency doubling the output from an amplified CPM dye laser. The change in transmission (ΔT) at a single wavelength was monitored by phase-sensitive detection of a continuum probe. The wavelength was tuned to the peak of the exciton and ΔT was recorded as a function of delay between pump and probe.

The exciton recovery time was investigated as a function of sample thickness, temperature, and pump intensity. The dependence on sample thickness at room temperature is shown in Fig. A.3. The recovery time (τ) was shorter for the thinner samples. It ranged from 110 ps for the 0.68 μm sample to 40 ps for the 0.23 μm sample, and probably results from enhanced surface recombination. τ was also seen to decrease as the temperature decreased. It decreased from 90 ps at room temperature to 50 ps at 150 K for the 0.55- μm thick sample. We tentatively attribute this to the increased exciton oscillator strength at lower temperatures, which leads to a shorter lifetime.

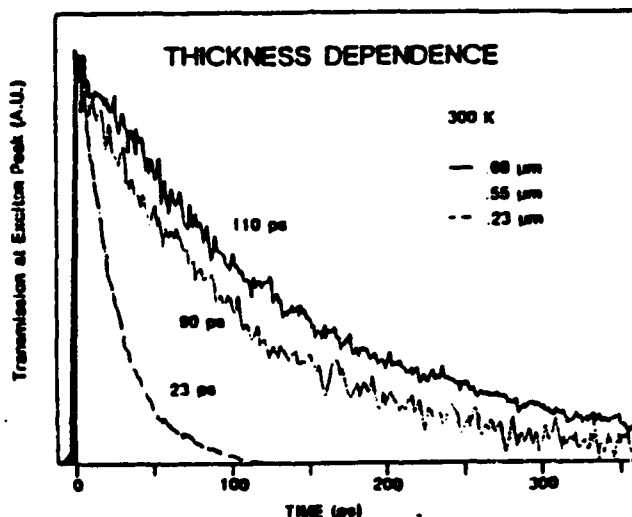


Figure A.3 Transmission at the exciton peak as a function of pump-probe delay for various thicknesses of ZnSe thin films at room temperature.

A.5 Conclusions

We observed the optical nonlinearities of electronic origin in MBE grown ZnSe thin films at room temperature and at low temperatures. The maximum value of the nonlinear index, $\Delta n/N$, is $\approx 1.8 \times 10^{-19} \text{ cm}^{-3}$ at room temperature and 9×10^{-19} at $T = 150 \text{ K}$. $\Delta n/N$ at room temperature is comparable to that obtained in bulk GaAs and GaAs/AlGaAs multiple quantum wells. The experimental results were compared with a partly phenomenological plasma theory that takes into account the many-body effects of exciton screening, band-gap renormalization, band filling, and broadening of the tail states. Good agreement between experiment and theory is obtained. Response times of the ZnSe thin films were also measured and values between 20–110 ps were obtained.

A.6 References

- A.1 F. V. Karpushko and G. V. Sinitzyn, *J. Appl. Spectrosc. USSR* 29, 1323 (1978).
- A.2 G. Olbright, N. Peyghambarian, H. M. Gibbs, A. Macleod, and F. Van Milligen, *Appl. Phys. Lett.* 45, 1031 (1984).
- A.3 S. D. Smith, J. G. H. Mathew, M. R. Taghizadeh, A. C. Walker, B. S. Wherrett, and A. Hendry, *Opt. Commun.* 51, 357 (1984).
- A.4 L. A. Kolodziejski, R. L. Gunshor, N. Otsuka, S. Datta, W. M. Becker, and A. V. Nurmikko, *IEEE J. Quantum Electron.* QE-22, 1666 (1986).
- A.5 S. H. Park, A. D. Jeffery, R. A. Morgan, S. W. Koch, N. Peyghambarian, J. E. Potts, and H. Cheng, *Digest of the CLEO Conference* (Optical Society of America, Washington, DC, 1987), post-deadline paper Thul3.
- A.6 N. Peyghambarian, S. H. Park, S. W. Koch, A. Jeffery, J. E. Potts, and H. Cheng, *Appl. Phys. Lett.* 52, 182 (1988).
- A.7 O. Madelung, M. Schulz, and H. Weiss, eds., *Landolt-Bornstein* (Springer-Verlag, Berlin, 1987).
- A.8 H. Cheng, S. K. Mohapatra, J. E. Potts, and T. L. Smith, *J. Cryst. Growth* 81, 512 (1987).
- A.9 J. E. Potts, H. Cheng, S. K. Mohapatra, and T. L. Smith, *J. Appl. Phys.* 61, 333 (1987).
- A.10 L. Banyai and S. W. Koch, *Z. Physik B* 63, 283 (1986); for a review of the different many-body effects, see also H. Haug, in *Optical Nonlinearities and Instabilities of Semiconductors*, H. Haug, ed. (Academic, New York, 1987); see also H. Haug and S. Schmitt-Rink, *Prog. Quantum Electron.* 9, 3 (1984).
- A.11 G. R. Olbright, N. Peyghambarian, S. W. Koch, and L. Banyai, *Opt. Lett.* 12, 413 (1987).
- A.12 Y. H. Lee, A. Chavez-Pirson, S. W. Koch, H. M. Gibbs, S. H. Park, J. F. Morhange, A. D. Jeffery, N. Peyghambarian, L. Banyai, A. C. Gossard, and W. Wiegmann, *Phys. Rev. Lett.* 57, 2446 (1986).

APPENDIX B: INTRODUCTION TO DICHROMATED GELATIN (DCG) RECORDING MATERIAL

B.1 Overview

The property of high-efficiency, low-distortion reconstruction makes volume phase holograms more attractive than other classes of holograms. To take advantage of this property, recording materials must have high spectral sensitivity and high refractive index modulation and must be transparent to reconstructing radiation.

Hardened, dichromated gelatin produces a near-optimum medium for holography. Holograms produced in this medium have the following properties:

1. High efficiency. Holograms with reflection efficiencies of up to 99.99% have been reported for a simple grating structure [B.1].
2. High signal-to-noise ratio, as scattering noise is minor.
3. High refractive index modulation, $\cong 0.01$ - 0.08 , with uniform spatial frequency response of over 100-5300 lines/mm [10.37].
4. Ability to obtain multiple exposures, each with diffraction efficiencies (DE) of $\cong 100\%$, on the same film.
5. A high damage threshold of 1 MW/cm², which indicates that DCG is good for high power operation [10.40].

Holograms in DCG also have some potential disadvantages:

1. Low photosensitivity, about 1000 times less sensitive than for photographic emulsions. In addition, the photosensitivity decreases strongly as the wavelength increases, limiting the exposure wavelength to the green or blue spectral range.
2. High dynamic change of Δn , which limits the number of hologram being recorded in the same material.
3. Strict requirements for processing emulsion, which makes readily usable, high quality films unavailable.
4. No clear photochemical mechanism to explain DCG behavior during processing [B.2]. The lack of such a mechanism has necessitated the exploration of many different development procedures.

In addition, there is increasing interest in the use of a DCG holographic grating for application in the red-to-near-IR regime. However, holograms recorded in the visible range but reconstructed in the near-IR range will require thicker gelatin

emulsions, as indicated by Eq. (10.26), if optical performance comparable to that of holograms reconstructed in the visible range is expected. The cast coating technique has been effective for obtaining gelatin emulsions of uniform thickness [B.3]. Other dye-sensitized techniques for DCG material may provide highly efficient gratings for near-IR applications [B.4].

B.2 Physical Properties of DCG

DCG is a gelatin sensitized with ammonium dichromate, $(\text{NH}_4)_2\text{Cr}_2\text{O}_7$, or with potassium dichromate, $\text{K}_2\text{Cr}_2\text{O}_7$. During exposure, light is absorbed to convert Cr^{+6} , possibly to Cr^{+3} , forming several intermediate chromium compounds. These compounds are insoluble in water; however, the gelatin, unexposed or inefficiently exposed, is soluble in water. The exposed film usually is developed in liquids that cause the emulsion to swell and that wash out the unreacted dichromate; the emulsion then is dehydrated to near its original thickness. Different exposure regions are hardened to different degrees, causing periodic strain (because of two-plane-wave interference, for example) during development. The exact mechanism by which the change in refraction index, Δn , takes place is a subject of intense debate. It is likely that air-filled voids of varying density are formed near the unhardened gelatin chains, resulting in a difference in refractive index between the gelatin and the voids. It is this difference which allows the formation of high-efficiency holograms.

Normal chemical processing of photographic emulsion usually results in shrinkage of the emulsion. The shrinkage may be attributable to varying quantities of liquid in the gelatin or to chemical effects such as hardening. In general, emulsion shrinkage will distort the recorded grating spacings, as in the case of multiple gratings. Each grating will have a different angular change, resulting in reduced DE and/or in distortion of the reconstructing beam. Moreover, emulsion shrinkage is far more significant for reflection-type holograms than for transmission-type holograms, as grating planes for reflection type are parallel to the emulsion surface. For more details on variations in the thickness of Kodak 649F caused by chemical processing, one can refer to Meyerhofer's work [B.2]. To reduce these undesirable effects in a transmission hologram, in recording the hologram it is best that the grating planes be normal to the emulsion surface ($\phi=\pi/2$), that is, it is best to use unslanted gratings.

Emulsion readily swells during processing, but all of the moisture (or excess chemicals) in the emulsion is not removed during dehydration with isopropanol. In addition, the amount of swelling is not uniform over the entire emulsion. For example, swelling in the center of film is 25% greater than that near the edges [B.5].

Therefore, reconstruction angles will not be the same over the entire emulsion. One solution to this problem is to bake the emulsion at an elevated temperature to drive out residual moisture. Unless baking temperature and baking time are optimized, however, DE will decrease with the reduction in grating thickness.

It appears that the refractive index of DCG depends on its water content, a property not fully under control because drying occurs continuously. However, it is believed that the refractive index modulation is caused by the rearrangement of gelatin molecule chains during development; this rearrangement in turn is the result of the hardness differential [10.37]. The refractive index change, Δn , in Eq. (10.26) can be described as

$$\Delta n = k_p k_{wd} k_f \Delta h \quad . \quad (B.1)$$

where $k_p \equiv (W - W_0)/W_0$, measured as a percentage increase in weight, a swelling factor after sensitization, with W_0 the weight of a dry film and W the weight of a swollen gelatin layer; k_{wd} is a swelling factor for water development; k_f is a constant gain factor (≈ 5) which takes into account the amplification of Δn in the dehydration process; and $\Delta h = (h_e - h_u)$ is the difference in hardness between exposed and unexposed regions.

B.3 Fabrication Procedures

B.3.1 Background

The fabrication procedure described below is based on the method developed by T. G. Georgekutty and H. K. Liu [B.6] whose work is similar to that of C. D. Leonard and B. D. Guenther [B.5]. Our experiments show different efficiency-to-exposure characteristic curves, however. Some development steps described in the procedure were derived solely from our experiments and coincidentally were very similar to those reported by Leonard and Guenther [B.7].

Our simplified method has the following advantages:

(a) Without index matching, a high DE of 86% at $\lambda = 514.5$ nm was consistently achieved with a moderate processing time ≈ 1.5 hour, compared to conventional methods which take ≈ 12 hour.

(b) The light energy density required for the exposure to achieve maximum efficiency is approximately 30–45 mJ/cm², about five times less than reported by Georgekutty and Liu [B.6] and by Lin [B.8].

(c) No noticeable difference was observed in DE for fabrication with tap water (pH = 6.97) as opposed to fabrication with distilled water.

(d) Lower susceptibility to relative humidity was observed.

Numerous trials over an extended period led to the following procedure. Further optimization to give a higher DE might be possible, for example, through index-matching techniques.

B.3.2 Fabrication of DCG HOEs with Kodak 649F Plates: A Simplified Method

1. Preprocessing (under room light)

- P1** Soak DCG sample in Kodak Rapid Fixer with standard hardener concentration (3.25%) for 10-15 min.
- P2** Wash in running water (20 C°-25 C°) for at least 15 min.
- P3** Soak in hot water at 63±2 C° for 10 min.

2. Sensitization (under dim red light)

- S1** Soak sample in 10% ammonium dichromate solution (with 1% Kodak Photo-Flo 200) in a flat pan with the emulsion side up for 5 min.
- S2** Remove residual ammonium dichromate from the glass side by first wet-wiping and then dry-wiping.
- S3** Bake at 71±10 C° for 5-10 min.; hold sensitized plates in darkness for 3-24 hrs.

3. Exposure

4. Development (under dim red light for steps D1 and D2, then under room light)

- D1** Soak sample in 0.5% solution of ammonium dichromate for 5 min.
- D2** Soak in Kodak Rapid Fixer with standard hardener concentration (3.25%) for 5 min.
- D3** Wash in running water for 10 min.
- D4** Dehydrate in a 50:50 or 60:40 isopropanol/distilled water solution for 3 min.; wipe off residuals from the glass side and edges.
- D5** Dehydrate in 100% isopropanol for 3 min., agitating vigorously (this step is critical).
- D6** Bake at 71±10 C° for 10 min. after wipe.

5. Assembly

- A1** Bake sample for 1 hour at 71±10 C°
- A2** Mix 5 cc. cement with one drop of catalyst.
- A3** Cement glass plates together and hold for 2 hours.

B.4 Notes of Interest for DCG Processing:

- * Experiments indicate that to ensure consistent results control of processing time and temperature is crucial. Soaking the sample in water of temperatures $>65^{\circ}\text{C}$ and/or for longer than 10 min. (step P3) cannot yield samples of sufficiently high DE, possibly because of dissociation of the gelatin emulsion.
- * The plate should be constantly agitated during preprocessing, sensitization, and development. Agitation helps ensure fabrication of holograms with less noise.
- * No significant difference in results was observed whether tap water or deionized water was used in soaking/washing steps.
- * The performance of Kodak 649F has been shown to vary according to batch and/or age.
- * Solutions should be discarded after one day of use (exceptions are the 50:50 alcohol/water solution and 100% isopropanol solution).
- * Holograms are insensitive to relative humidities of up to 70-80% after proper baking, which demonstrates good environmental stability.
- * A good way to protect holograms from moisture penetration is to cement a cover glass to the hologram with a transparent epoxy or an optical glue (e.g., Norland 11). No noticeable drop in DE was observed after cementing.
- * If plates are exposed too soon after sensitization, extra scattering noise in the developed hologram is observed. If the sensitized plates are held too long before exposure, they begin to lose sensitivity. A tradeoff is to keep DCG in darkness for about 3-24 hours after preprocessing. Prehardness, hence the refractive index modulation, will be increased due to dark reaction.
- * The correct exposure for maximum DE depends on the percentage of refractive index modulation required and on the concentration of ammonium dichromate. Test exposures should be run to ensure optimum performance. We suggest a 50 mJ/cm^2 exposure and an ammonium dichromate concentration of 5%.
- * The concentration of ammonium dichromate controls the exposure time and the Bragg angle shift; the amount of hardener controls the DE and scattering noise.
- * If a plate is milky white in appearance after development, it could be that the solution concentration at step D1 and/or the drying time at step S3 are

not appropriate, and/or the hardening time at step P1 is insufficient. However, if the drying temperature is too high, the gelatin can be scorched and become opaque; this occurrence is easily observed under dim red light in subsequent steps.

- * Multiple internal reflections inside the emulsion will interfere with incident beams during the recording process, creating unwanted secondary holograms which are quite stable and which can decrease the DE by up to 50%. An index-matching technique can be applied to reduce this effect.

B.5 Safety Guide

Ammonium dichromate has the form of bright orange crystals or power. It is odorless and stable at room temperature. Ammonium dichromate is a strong oxidizer and can react violently with easily oxidizable substances such as paper, wood, and textiles. Prolonged or repeated inhalation of dichromate dust may result in ulceration and perforation of the nasal septum. Although ammonium dichromate has been identified as a "noncarcinogenic chromium", chromium compounds are assumed to be carcinogenic to humans.

Following are important health considerations for persons working with with ammonium dichromate.

1. Avoid generation of dust.
2. Use protective equipment to prevent skin and eye contact and to prevent inhalation of ammonium dichromate dust.
3. Ammonium dichromate solution may require disposal as hazardous waste. Call risk management department for help.
4. General and local exhaust ventilation is required.

B.6 References

- B.1 T. Jansson and J. Jansson, Technical Report No. 84-1479 (1984).
- B.2 D. Meyerhofer, RCA Review 33, 110 (1972).
- B.3 S. S. Duncan, J. A. McQuoid, and D. J. Cartney, Proc. SPIE 523, 196 (1985).
- B.4 R. Changkakoti and S. V. Pappu, Appl. Opt. 25, 798 (1986).
- B.5 T. A. Shankoff, Appl. Opt. 7, 2101 (1968).
- B.6 T. G. Georgekutty and H.-K. Liu, Appl. Opt. 26, 371 (1987).
- B.7 C. D. Leonard and B. D. Guenther, Technical Report T-79-17 (1979).
- B.8 L. H. Lin, Appl. Opt. 8, 963 (1969).

APPENDIX C: Conclusions Section of the 1988 PhD dissertation of George Gigioli, Jr. entitled "Optimization and Tolerancing of Nonlinear Fabry-Perot Etalons for Optical Computing"

[Gigioli treats a circuit consisting of identical gates for each of four modes of logic gate operation: transmission mode threshold logic (TMTL), transmission mode latching logic (TMLL), reflection mode threshold logic (RMTL), and reflection mode latching logic (RMLL). He assumes a Kerr nonlinearity. His pessimistic conclusions about steady-state, one-wavelength logic would probably be the same for the carrier-density-dependent nonlinearity of GaAs. The conclusions are much less clear for NOR-gate operation of a GaAs etalon.]

What started out to be the groundwork for a systematic approach to the design of a nonlinear Fabry-Perot etalon for use as a digital optical switching device turned out to be a denial of its feasibility. While the analysis falls short of a formal proof, the conclusions are well supported by the evidence presented in the examples. In this final chapter, this evidence is recapitulated through a summary of the conclusions reached in the other chapters, and the means by which they were reached. Because of the negative nature of the results obtained, care is taken here to reiterate the limits of validity of the analysis, so that the conclusions are not inappropriately assumed in areas where they do not apply.

In Chapter Three it was shown how a mathematical model of a nonlinear etalon can form the basis of an optimization procedure for the etalon. Inasmuch as the model accurately predicts its behavior, a nonlinear etalon can be optimized with respect to several of its transfer function parameters. The procedure involves constructing a merit function in which each parameter is represented mathematically by a term in the function. An optimized design is one for which the merit function is maximized. Since each of the parameters pushes this maximum in a different direction, the optimization of the etalon becomes a tradeoff among the optimizations of individual characteristics of the device. The relative importance of these characteristics to the overall design is the basis for choosing the weightings of the terms in the merit function (they obviously needn't all be weighted equally).

The particular design which is chosen as a result of this optimization really depends on how the terms of the merit function are constructed. In the example merit functions of Chapter Three, this translates into the acceptance limits that were placed on the device characteristics (e.g., maximum threshold intensity, minimum contrast, etc.), and the choice of parameter for which the initial detuning was chosen. In the example, the detuning was chosen to maximize the contrast, but it could just

as well have been chosen to minimize the threshold intensity, which would have yielded a different design. Other important considerations in the optimization are the design constraints. If the optimization is overconstrained, the result will be ambiguous, as was shown for the transmission-mode case when αL was equal to 0.3.

Because the merit function terms are based on the mathematical model of the etalon, the optimization strictly applies only to devices that are described by the equations of Chapter Two. Some of these include ZnS and ZnSe interference filters, and semiconductor devices operating far from any electronic resonances. It does not apply to devices operating near resonance, in which the nonlinearity is a result of interaction with an exciton, band filling effects, or increasing absorption effects. In such devices, there are additional degrees of freedom which have not been included in this model, and which may lead to different solutions and different device tolerances.

It should be noted that the merit function maxima found in Chapter Three are not global maxima, since only one parameter of the device is used to pick the optimum initial detuning. The initial detuning could have been chosen with all performance parameters in mind; however, it was felt that, from the point of view of device performance, the contrast was the single most important parameter. For this reason, the detuning at each point in R_F - R_B space was chosen to maximize this. [R_F and R_B are the reflectivities of the front and back mirrors.] The other parameters were used to locate the merit function maximum with respect to the mirror reflectances.

In Chapter Four, it was found that, regardless of which logic mode is chosen, the transfer function tolerances are quite small—on the order of 1 to 1-1/2%, depending on the fan in and fan out assumed. That one type of logic mode has no particular advantage over any other, as far as tolerances are concerned, suggests that the tight tolerances are not a result of the particular shapes of the switching functions but rather of the fact that the signal contrast is low and that signal errors propagate. Based on this, it would seem not to matter whether or not the global maximum of the merit function were found since the tolerances would not be greatly improved.

Unlike the optimization procedure, the tolerance analysis can be applied to other types of devices, since it does not depend on the device model given in Chapter Two. All that is necessary is that the transfer function be approximated by two straight lines. Of course, in order to translate the resulting transfer function tolerances into cavity-parameter tolerances, some model of the device response must be used.

The major problem with the analysis of Chapter Four is that the approximations are weak. The actual transfer functions depart significantly from the straight-line approximations, yet the tolerances derived are around 1%. Unfortunately, approximating the curves more closely by using higher-order functions further complicates the analysis to the point of intractability. The equations are complicated enough with linear approximations. Besides, it is reasonable to suspect that closer approximations will only make the tolerances tighter since, in a sense, some of the tolerance is used up by the nonlinearity of the approximating curve. In this light, the analysis seems optimistic.

In Chapter Five, the transfer-function tolerances were translated back into cavity-parameter tolerances using the model of Chapter Two. It was found that the cavity-parameter tolerances are unrealistic for large-scale fabrication of gates. One may then conclude that digital optical computing with nonlinear Fabry-Perot etalons is unfeasible. All four logic modes are about equally unfavorable, with reflection mode latching logic being slightly worse than the others. This is largely due to the narrow bistable region of the transfer curve which, in effect, squeezes in the tolerances.

The estimates of errors in the transfer function due to perturbations in the cavity parameters are subject to the linear approximations to the transfer functions, so the derived parameter tolerances are really only within an order of magnitude of the true tolerances. Nonetheless, this order of magnitude is quite small compared to the tolerances of around 10%, that would make the devices producible. If they had been closer, the results may have been disputed, based on the level of approximation. The error budget allocates equal error contributions to each of the parameters; however, since some parameters tend to vary more than others, they require proportionally larger tolerances. The tolerances may be improved slightly by allowing more sensitive parameters to have a wider variation. Again, this is not expected to make a large improvement in the tolerances of the devices.

In general, even though some of the approximations are coarse, and some assumptions questionable, the final conclusion regarding the feasibility is valid for these types of devices. A similar analysis should be performed for devices that do not follow the model of Chapter Two but that are being proposed for use as digital optical logic elements. Some of these (such as SEEDs and two-wavelength NOR-gate etalons) will probably not be as sensitive to detuning variations as the devices studied here; however, no estimates on the tolerances of these devices are conjectured here.

Optically bistable devices which operate near an electronic resonance are not described by the model of Chapter Two, so the optimization procedure of Chapter

Three does not apply to them. In particular, since the nonlinear index and the absorption both depend on the internal intensity and are not described simply by an n_2 -type coefficient, there are more degrees of freedom with which to influence the optimization. On the other hand, since both the nonlinear index and absorption are dependent on the detuning of the radiation frequency from the electronic resonance (not to be confused with the detuning of the radiation from the cavity resonance or initial detuning as it has been referred to in this dissertation), immediately there are more material parameters to account for in the tolerance analysis. For this reason alone, it is suspected that the tolerances for these types of devices will not prove to be any more favorable than the ones analyzed here.

Nonlinear etalon devices may still be useful for other types of computing applications such as analog systems and neural networks in which the etalon is not operating in bistable mode and in which errors in the transfer curves are more tolerable. Also, they may be useful in digital systems that require only a small number of threshold devices with the bulk of the computation being handled by appropriate interconnections between gates ("wired OR" logic), or in hybrid systems in which signals might be converted back and forth between optical and electronic so that signal errors do not propagate. Hopefully, the optimization procedure developed herein can be applied to these types of device applications.

In closing, the results presented here may explain the scarcity in the literature of descriptions or demonstrations of medium- to large-scale switching circuits based on nonlinear Fabry-Perot etalon gates. It is surprising, however, that an analysis such as this had not been performed earlier. Hopefully, these results will prompt others to perform similar studies on other types of optical switching devices in order to determine whether large-scale digital computing systems are possible before trying to build them and make them work.

DISTRIBUTION LIST

addresses	number of copies
Dorothy J. Jackson RADC/DCLW	20
RADC/DOVL GRIFFISS AFB NY 13441	1
RADC/IMP GRIFFISS AFB NY 13441	2
ADMINISTRATOR DEF TECH INF CTR DTIC-FDA) CAMERON STA BG 5 ALEXANDRIA VA 22304-6145	5
RADC/COTD BLDG 3, ROOM 16 GRIFFISS AFB NY 13441-5700	1
AFCSA/SAMI Attn: Miss Griffin 10363 Pentagon Wash DC 20330-5425	1
HQ USAF/SCIT Pentagon Wash DC 20330-5190	1
SAF/AQSC Pentagon 4D-267 Wash DC 20330-1000	1
DIRECTOR DMAHTC ATTN: SDSIM Wash DC 20315-0030	1

Director, Info Systems 1
OASD (C3I)
Rm 3E187
Pentagon
Wash DC 20301-3040

Fleet Analysis Center 1
Attn: GIDEP Operations Center
Code 3061 (E. Richards)
Corona CA 91720

HQ AFSC/XTKT 1
ANDREWS AFB DC 20334-5000

HQ AFSC/XRK 1
ANDREWS AFB MD 20334-500

HQ SAC/SCPT 1
OFFUTT AFB NE 68113-5001

DTESA/ROE 1
ATTN: LARRY G. MCMANUS
KIRTLAND AFB NM 87117-5000

HQ TAC/DRIY 1
Attn: Mr. Westerman
Langley AFB VA 23665-5001

HQ TAC/DRCA 1
LANGLEY AFB VA 23665-5001

ASD/ENEMS 2
Wright-Patterson AFB OH 45433-6503

ASD/AFALC/AXAE 1
Attn: W. H. Dungey
Wright-Patterson AFB OH 45433-6533

WRDC/AAAI WRIGHT-PATTERSON AFB OH 45433-6533	1
AFIT/LDEE BUILDING 640, AREA B WRIGHT-PATTERSON AFB OH 45433-6583	1
WRDC/MLFO WRIGHT-PATTERSON AFB OH 45433-6533	1
AAMRL/HE WRIGHT-PATTERSON AFB OH 45433-6573	1
Air Force Human Resources Laboratory Technical Documents Center AFHRL/LRS-TDC Wright-Patterson AFB OH 45433	1
2750 ABW/SSLT Blds 262 Post 11S Wright-Patterson AFB OH 454433	1
AUL/LSE MAXWELL AFB AL 36112-5564	1
Defense Communications Engineering Ctr Technical Library 1860 Wiehle Avenue Reston VA 22090-5500	1
COMMAND CONTROL AND COMMUNICATIONS DIV DEVELOPMENT CENTER MARINE CORPS DEVELOPMENT & EDUCATION COMMAND ATTN: CODE DIOA QUANTICO VA 22134-5080	2

AFMCM/UDY
ATTN: CH, SYS ENGR DIV
GUNTER AFS AL 36114

1

U.S. Army Strategic Defense Command
Attn: DASD-H-MPL
P.O. Box 1500
Huntsville AL 35807-3801

1

COMMANDING OFFICER
NAVAL AVIONICS CENTER
LIBRARY - D/765
INDIANAPOLIS IN 46219-2189

1

COMMANDING OFFICER
NAVAL TRAINING SYSTEMS CENTER
TECHNICAL INFORMATION CENTER
BUILDING 2068
ORLANDO FL 32813-7100

1

COMMANDER
NAVAL OCEAN SYSTEMS CENTER
ATTN: TECHNICAL LIBRARY, CODE 9642B
SAN DIEGO CA 92152-5000

1

COMMANDER (CODE 3433)
ATTN: TECHNICAL LIBRARY
NAVAL WEAPONS CENTER
CHINA LAKE, CALIFORNIA 93555-6001

1

SUPERINTENDENT (CODE 1424)
NAVLA POST GRADUATE SCHOOL
MONTEREY CA 93943-5000

1

COMMANDING OFFICER
NAVAL RESEARCH LABORATORY
ATTN: CODE 2627
WASHINGTON DC 20375-5000

2

SPACE & NAVAL WARFARE SYSTEMS COMMAND
PMW 153-3DP
ATTN: R. SAVARESE
WASHINGTON DC 20363-5100

1

CDR, U.S. ARMY MISSILE COMMAND
REDSTONE SCIENTIFIC INFORMATION CENTER
ATTN: AMSMI-RD-CS-R (DOCUMENTS)
REDSTONE ARSENAL AL 35898-5241

2

Advisory Group on Electron Devices 2
Hammond John/Technical Info Coordinator
201 Varick Street, Suite 1140
New York NY 10014

UNIVERSITY OF CALIFORNIA/LOS ALAMOS 1
NATIONAL LABORATORY
ATTN: DAN BACA/REPORT LIBRARIAN
P.O. BOX 1663, MS-F364
LOS ALAMOS NM 87545

SAND CORPORATION THE/LIBRARY 1
HELPER DORIS S/HEAD TECH SVCS
P.O. BOX 2138
SANTA MONICA CA 90406-2138

AEDC LIBRARY (TECH REPORTS FILE) 1
MS-100
ARNOLD AFS TN 37389-9998

USAG 1
Attn: ASH-PCA-CRT
Ft Huachuca AZ 85613-6000

DOT LIBRARY/10A SECTION 1
ATTN: M493.2
800 INDEPENDENCE AVE. S.W.
WASH DC 20591

1839 EIG/EIET (KENNETH W. IRBY) 1
KEESLER AFB MS 39534-6348

JTFPO-TD 1
Dr. Raymond F. Freeman
Director, Advanced Technology
1500 Planning Research Drive
McLean, Virginia 22102

HQ ESC/CWPP 1
San Antonio TX 78243-5000

AFEWC/ESRI 3
SAN ANTONIO TX 78243-5000

485EIG/EIR/M Craft
GRIFFISS AFB NY 13441-6348

1

ESD/XTP
ATTN: ADV SYS DEV
HANSCOM AFB MA 01731-5000

1

ESD/ICP
HANSCOM AFB MA 01731-5000

1

ESD/AVSE
BLDG 1704
HANSCOM AFB MA 01731-5000

2

HQ ESD SYS-2
HANSCOM AFB MA 01731-5000

1

The Software Engineering Institute
Attn: Major Dan Burton, USAF
Joint Program Office
Carnegie Mellon University
Pittsburgh PA 15213-3890

1

DIRECTOR
NSA/CSS
ATTN: T513/TDL (DAVID MARJARUM)
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: W166
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R-8316 (MR. ALLEY)
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R24
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R21
9800 SAVAGE ROAD
FORT GEORGE G MEASDE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: DEFSMAC (JAMES E. HILLMAN)
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R31
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R5
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R8
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: S21
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: W07
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: W3
FORT GEORGE G MEADE MD 20755-6000

1

DIRECTOR
NSA/CSS
ATTN: R523
FORT GEORGE G MEADE MD 20755-6000

2

DoD COMPUTER SECURITY CENTER
ATTN: C4/TIC
9900 SAVAGE ROAD
FORT GEORGE G MEADE MD 20755-6000

1

Heatt M. Gibbs
University of Arizona
Optical Sciences Center
Tucson, Arizona 85721

5

DirectorDARPA
1400 Wilson Blvd
Arlington, Va 22209-2308

5